

Result No.	Query			Length	DB	ID	Description
	Score	Match					
1	696	100.0	508	4	AAU04693	Human suc	
2	696	100.0	508	4	ADC35101	Human bre	
3	696	100.0	521	4	AB111822	Human FGF	
4	685	98.4	508	4	AAW62558	Fibroblas	
5	391	56.2	101	2	AAV07018	Breast ca	
6	174	25.0	442	4	AB368120	Drosophi	
7	143.5	20.6	481	2	AAW62507	p62 prote	
8	143.5	20.6	509	7	ADD14099	Human src	
9	143.5	20.6	541	5	ABG66871	Human pro	
10	141	20.3	359	3	AAV79184	Haematopo	
11	137.5	19.8	331	6	ABR83336	Human NOV	
12	137.5	19.8	337	4	AAW39806	Human pol	
13	137.5	19.8	365	4	AAW41592	Human pol	
14	137.5	19.8	365	4	AAW41593	Human pol	
15	133	19.4	412	3	AB52062	Gene 2 hu	
16	134.5	19.3	622	4	ABW58836	Drosophi	
17	133.5	19.2	352	4	AAW39807	Human pol	
18	127.5	18.3	315	4	AAW40528	Human pol	
19	124.5	17.9	306	4	AAW38742	Human pol	
20	124.5	17.9	326	4	AAW39794	Human pol	
21	124.5	17.9	326	4	AB592876	Human pro	
22	124.5	17.9	338	7	AD071132	Human int	
23	124.5	17.9	371	4	AAW41580	Human pol	
24	123.5	17.7	365	4	AAW37303	Human gen	
25	114.5	16.5	473	4	ABG18438	Novel hum	

PI Zhou M;
 XX WPI; 2001-425868/45.
 XX New isolated nucleic acid sequences encoding polypeptides useful in
 PT structure based rational drug design.
 XX
 XX Claim 7; Page 228-230; 235pp; English.
 PS
 XX The sequence represents the amino acid sequence of human suc-1-associated
 CC neurotrophic factor target protein (SNT-1). The invention provides
 CC fragments of SNT and fibroblast growth factor receptor (FGFR) which can
 CC form a binding complex that is amenable to structural determinations by
 CC nuclear magnetic resonance (NMR) spectroscopy. The invention provides
 CC methodology for related structure based rational drug design using the
 CC three-dimensional data. The invention fulfills the need for the
 CC determination of the three-dimensional structure of such complexes as
 CC described and for procedures for related structural base drug design
 CC predicated on such structural data
 XX
 XX Sequence 508 AA;
 SQ
 Query Match 100.0%; Score 696; DB 4; Length 508;
 Best Local Similarity 100.0%; Pred. No. 8e-75;
 Matches 130; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 DTVPDNRHNFVKVINVDGNGELSGIMELTDTLLIYTRKRSVKWHYLCRLRYGYDSN 60
 DB 11 DTVPDNRHNFVKVINVDGNGELSGIMELTDTLLIYTRKRSVKWHYLCRLRYGYDSN 70
 QY 61 LFSFSGRRCTGGGIFAFKCARAEELFNMLQEIOMNNSINVVEEPPVVERNNHOTELEVP 120
 DB 71 LFSFSGRRCTGGGIFAFKCARAEELFNMLQEIOMNNSINVVEEPPVVERNNHOTELEVP 130
 QY 121 RTPRTPTTPG 130
 DB 131 RTPRTPTTPG 140
 RESULT 2
 ADC35101
 ID ADC35101 standard; protein; 508 AA.
 AC ADC35101;
 DT 18-DEC-2003 (first entry)
 DE Human breast cancer antigen seq id 67.
 KW breast cancer; breast cancer diagnosis; breast cancer antigen.
 OS Homo sapiens.
 XX US2003108888-A1.
 XX
 XX 12-JUN-2003.
 XX
 XX 15-MAY-2002; 2002US-00146473.
 XX
 XX 15-MAY-2001; 2001US-0291150P.
 XX (LUDW-) LUDWIG INST CANCER RES.
 XX
 XX Scanlan MJ, Gout I, Stockert E, Old LJ, Gure A, Chen Y;
 XX WPI; 2003-829397/77.
 XX N-PSDB; ADC35143.
 XX
 XX Diagnosing breast cancer in subject by obtaining biological sample from
 PT subject, contacting sample with breast cancer-associated polypeptides,
 PT determining specific binding between polypeptides and agents in sample.
 XX
 XX Example 2; SEQ ID NO 67; 173pp; English.
 PS

XX The invention describes a method of diagnosing breast cancer in subject
 CC comprising contacting biological sample from subject with at least two
 CC different breast cancer-associated polypeptides (I) encoded by nucleic
 CC acid molecules (II) comprising sequence chosen from 42 fully defined
 CC sequences as given in specification, determining specific binding between
 CC (I) and agents in sample, where presence of the binding is diagnostic for
 CC breast cancer. The method is useful for diagnosing breast cancer in a
 CC subject. The sample is blood, lymph node fluid or breast discharge fluid.
 CC This is the amino acid sequence of a breast cancer antigen.
 XX
 XX Sequence 508 AA;
 SQ
 Query Match 100.0%; Score 696; DB 7; Length 508;
 Best Local Similarity 100.0%; Pred. No. 8e-75;
 Matches 130; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 DTVPDNRHNFVKVINVDGNGELSGIMELTDTLLIYTRKRSVKWHYLCRLRYGYDSN 60
 DB 11 DTVPDNRHNFVKVINVDGNGELSGIMELTDTLLIYTRKRSVKWHYLCRLRYGYDSN 70
 QY 61 LFSFSGRRCTGGGIFAFKCARAEELFNMLQEIOMNNSINVVEEPPVVERNNHOTELEVP 120
 DB 71 LFSFSGRRCTGGGIFAFKCARAEELFNMLQEIOMNNSINVVEEPPVVERNNHOTELEVP 130
 QY 121 RTPRTPTTPG 130
 DB 131 RTPRTPTTPG 140
 RESULT 3
 ABB11822
 ID ABB11822 standard; peptide; 521 AA.
 XX
 AC ABB11822;
 XX
 XX 11-JAN-2002 (first entry)
 DT
 DE Human FGFR signalling adaptor SNT-1 homologue, SEQ ID NO:2192.
 XX
 KW Human; cytokine; cell proliferation; cell differentiation; growth factor;
 KW haematopoiesis regulation; tissue growth; immunomodulator; activin;
 KW inhibin; chemotaxis; chemokinesis; thrombolysis; oncogenesis;
 KW proliferation; metastasis; cancer; tumour; haematopoietic disorder;
 KW myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;
 KW chronic inflammatory condition; proliferative retinopathy;
 KW atherosclerosis; coronary heart disease; arterial ischaemia;
 KW bone disorder; osteoporosis; vascular growth disorder;
 KW tissue regeneration; wound healing; infection; immune disorder;
 KW cell culture; drug screening; gene therapy; anti-inflammatory;
 KW antasthmatic; antiarthritic; haemostatic; antiarteriosclerotic;
 KW cytostatic; osteopathic; vasotropic; cardiant; virucide; antibacterial;
 KW antifungal; vulnery; antiulcer.
 XX
 OS Homo sapiens.
 XX
 XX WO200157188-A2.
 XX
 XX 09-AUG-2001.
 XX
 XX 05-FEB-2001; 2001WO-US003800.
 XX
 XX 03-FEB-2000; 2000US-00496914.
 XX 27-APR-2000; 2000US-00560875.
 XX (HYSE-) HYSEQ INC.
 XX
 XX Tang YT, Liu C, Drmanac RT;
 XX WPI; 2001-457740/49.
 XX N-PSDB; ABA09066.
 XX
 XX Human proteins and DNA encoding sequences useful for preventing, treating
 PT

PT or ameliorating a medical condition in a mammalian subject e.g. arthritis
 PT and cancer.

Claim 20; Page 261; 1963pp; English.

Sequences ABB10981-ABB12330 represent 1350 novel human polypeptides, and sequences ABA08225-ABA09574 represent nucleic acids encoding them. The invention also relates to vectors and recombinant host cells comprising a nucleotide of the invention, methods of producing the novel polypeptides, antibodies against the polypeptides, methods of detecting the nucleotides or polypeptides in a sample, and methods of identifying compounds which bind to polypeptides of the invention. Although novel, many of the polypeptides of the invention have homology to known proteins, thereby giving an insight into their probable biological activities, and hence potential therapeutic applications. The polypeptides of the invention may have various activities, including cytokine, cell proliferation or cell differentiation activities; stem cell growth factor activity; haematopoiesis regulatory activity; tissue growth activity; immunomodulatory activity; activin- or inhibin-related activities; chemotactic or chemokinetic activities; haemostatic, thrombotic or thrombolytic activities; receptor or ligand activities; or may be involved in oncogenesis, cancer cell proliferation or metastasis. Depending on their biological activities, polypeptides and nucleotides of the invention are useful for preventing, treating or ameliorating medical conditions, e.g., by protein or gene therapy. Such conditions include cancers, haematopoietic disorders (e.g., myeloid or lymphoid cell disorders), chronic inflammatory conditions (e.g., asthma or arthritis), proliferative retinopathy, atherosclerosis, coronary heart disease, arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal vascular growth. Polypeptides involved with tissue regeneration and repair (or nucleic acids encoding them) may be used to promote wound healing (e.g., of burns, incisions and ulcers), while those with immunomodulatory activities may be used in the treatment of viral, bacterial and fungal infections in addition to immune disorders. Polypeptides with growth factor activity may be used in cell cultures to promote cell growth. For example, such polypeptides may be used to manipulate stem cells in culture to give rise to neuroepithelial cells that can be used to augment or replace cells damaged by illness, autoimmune disease or accidental damage. The polypeptides and nucleotides may also be used in the diagnosis of the above conditions, and in drug screening techniques. The present sequence represents a novel human polypeptide of the invention

Sequence 521 AA;

Query Match 100.0%; Score 696; DB 4; Length 521;
 Best Local Similarity 100.0%; Pred. No. 8.3e-75;
 Matches 130; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DTPVDNHRNKFVINVDGNGELGSGIMELTDTEILYTRKDSVKWHYLCRLRYGYDSN 60
 DB 24 DTPVDNHRNKFVINVDGNGELGSGIMELTDTEILYTRKDSVKWHYLCRLRYGYDSN 83
 QY 61 LFSFESGRCCTGGGIFAFKCARAEELFNMLQEIQMNNNSINVVEPVPVVRNNHOTELEVP 120
 DB 84 LFSFESGRCCTGGGIFAFKCARAEELFNMLQEIQMNNNSINVVEPVPVVRNNHOTELEVP 143
 QY 121 RTRPTPTTPG 130
 DB 144 RTRPTPTTPG 153

RESULT 4

AAW62558
 ID AAW62558 standard; peptide; 508 AA.

AC AAW62558;

DT 25-MAR-2003 (revised)

XX 25-SEP-1998 (first entry)

DE Fibroblast growth factor receptor protein kinase substrate 2 (FRS2).

XX

KW Adapter protein; regulation; growth factor stimulation;
 KW fibroblast growth factor receptor protein kinase substrate 2; FRS2;
 KW cellular differentiation; cellular proliferation; stimulated fibroblast growth factor receptor; Ras/MAPK cascade;
 KW Grb-2/Sos complex; diagnosis; prevention; treatment.

OS Homo sapiens.

Key Location/Qualifiers
 Domain 11..152
 /note="phosphotyrosine binding domain"

PN WC9824902-A1.

XX 11-JUN-1998.

XX 01-DEC-1997; 97WO-US021851.

XX 03-DEC-1996; 96US-0032093P.

XX (SUGS-) SUGEN INC.

PA (UTNY-) UNIV NEW YORK MEDICAL CENT.

XX Kouhara H, Spivakkroizman T, Lax I, Schlessinger J;

XX WPI; 1998-333322/29.

XX New human adaptor protein FRS2 - useful in, e.g. diagnosis, prevention or treatment of abnormal conditions associated with cell proliferation or cell differentiation.

PS Claim 3; Fig 1A; 53pp; English.

XX The present sequence represents an adaptor protein linking protein kinases to activating partners in cells. The protein is designated CC fibroblast growth factor receptor protein kinase substrate 2 (FRS2). It CC regulates the growth factor stimulation of cellular differentiation and CC cellular proliferation by linking stimulated fibroblast growth factor CC receptor to the Ras/MAPK cascade via Grb-2/Sos complex. The FRS2 protein CC is useful to diagnose abnormal conditions associated with cell CC proliferation or cell differentiation in organisms caused by an CC aberration in a signal transduction pathway characterised by an CC interaction between a FRS2 polypeptide and a natural binding partner CC (especially in mammals). Such abnormal conditions can be associated with CC cell proliferation or cell differentiation. The protein is also useful to CC identify compounds that enhance or block FRS2 interactions with natural CC binding partners, to diagnose, prevent or treat abnormal conditions. CC (Updated on 25-MAR-2003 to correct PA field.) (Updated on 25-MAR-2003 to CC correct PI field.)

XX Sequence 508 AA;

Query Match 98.4%; Score 685; DB 2; Length 508;
 Best Local Similarity 97.7%; Pred. No. 1.7e-73;
 Matches 127; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 DTPVDNHRNKFVINVDGNGELGSGIMELTDTEILYTRKDSVKWHYLCRLRYGYDSN 60
 DB 11 DTPVDNHRNKFVINVDGNGELGSGIMELTDTEILYTRKDSVKWHYLCRLRYGYDSN 70
 QY 61 LFSFESGRCCTGGGIFAFKCARAEELFNMLQEIQMNNNSINVVEPVPVVRNNHOTELEVP 120
 DB 71 LFSFESGRCCTGGGIFAFKCARAEELFNMLQEIQMNNNSINVVEPVPVVRNNHOTELEVP 130
 QY 121 RTRPTPTTPG 130
 DB 131 RTRPTPTTPG 140

RESULT 5

AAV07018

ID AAV07018 standard; protein; 101 AA.

XX

AC AAY07018;
XX
DT 02-JUL-1999 (first entry)
DE
XX
DE Breast cancer associated antigen precursor sequence.
XX
XX Cancer associated antigen; diagnosis; research; treatment; human;
KW breast cancer; colon cancer; gastric cancer; renal cancer; lung cancer;
KW prostate cancer.
XX
XX Homo sapiens.
OS
XX WO9904265-A2.
FN
XX
PD 28-JAN-1999.
XX
PF 15-JUL-1998; 98MO-US014679.
XX
PR 17-JUL-1997; 97US-00896164.
PR 10-OCT-1997; 97US-0061599P.
PR 10-OCT-1997; 97US-0061765P.
PR 10-OCT-1997; 97US-00948705.
PR 11-OCT-1997; 97GB-00021697.
PR 22-JUN-1998; 98US-00102322.
XX
XX (LUDW-) LUDWIG INST CANCER/RES.
PA
XX Old LJ, Scanlan MJ, Stockert E, Gure A, Chen Y, Gout I;
PI O'hare M, Obata Y, Pfreundschuh M, Tureci O, Sahin U;
PI
XX WPI; 1999-132448/11.
DR
XX
XX New isolated cancer associated nucleic acids and polypeptides - isolated
PT using sera from cancer patients, used to develop products for the
PT diagnosis, monitoring or treatment of cancers.
PT
XX
PS Disclosure; Page 398; 787pp; English.
XX
XX The invention relates to a method for diagnosing a disorder characterised
CC by expression of a human cancer associated antigen precursor coded for by
CC a nucleic acid molecule (NAM). The method comprises: (a) contacting a
CC biological sample isolated from a subject with an agent that specifically
CC binds to the NAM, an expression product or a fragment of an expression
CC product complexed with an HLA molecule; and (b) determining the
CC interaction between the agent and the NAM or the expression product as a
CC determination of the disorder. The products and methods can be used in
CC the diagnosis, monitoring, research, or treatment of conditions
CC characterised by the expression of various cancer associated antigens.
CC The invention provides nucleic acid sequences and encoded polypeptides
CC which are cancer associated antigen precursors expressed in human breast
CC cancer, renal cancer, colon cancer, gastric cancer, prostate cancer and
CC lung cancer
XX
SQ Sequence 101 AA;
Query Match 56.2%; Score 391; DB 2; Length 101;
Best Local Similarity 98.6%; Pred. NO. 5.9e-39;
Matches 72; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 DTVPDHNKFKVINVDGNGELSGIMELDTLILYTRKDSVKWHYLCRLRYGYDSN 60
DB 24 DTVPDHNKFKVINVDGNGELSGIMELDTLILYTRKDSVKWHYLCRLRYGYDSN 83
QY 61 LFSFSGRRCTG 73
DB 84 LFSFSGRRCTG 96
RESULT 6
ABB68120
ID ABB68120 standard; protein; 442 AA.
XX
AC ABB68120;

XX 26-MAR-2002 (first entry)
DT
XX Drosophila melanogaster polypeptide SEQ ID NO 31152.
DE
XX Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.
KW
XX Drosophila melanogaster.
OS
XX WO200171042-A2.
PN
XX 27-SEP-2001.
PD
XX 23-MAR-2001; 2001WO-US009231.
PF
XX 23-MAR-2000; 2000US-0191637P.
PR 11-JUL-2000; 2000US-00614150.
XX (PEXE) PE CORP NY.
PA
XX Venter JC, Adams M, Li PWD, Myers EW;
PI
XX WPI; 2001-656860/75.
DR N-PSDB; ABL12223.
DR
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signaling and cell-cell
PT interactions.
PT
XX Disclosure; SEQ ID NO 31152; 21pp + Sequence Listing; English.
PS
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABJ01840-ABL16175) and the encoded proteins (ABBS7737-
CC ABB72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 442 AA;
Query Match 25.0%; Score 174; DB 4; Length 442;
Best Local Similarity 34.4%; Pred. NO. 7.6e-12;
Matches 42; Conservative 22; Mismatches 52; Indels 6; Gaps 2;
QY 7 HNKPKV----INVDDGNGELSGIMELDTLILYTRKDSVKWHYLCRLRYGYDSNLF 62
DB 16 HENVFRVVAHQPTGPTPIKSGYELTPRELIFQSPGCEPIWVALQHLARYGLNLDLF 75
QY 63 SFESGRRCTGGGIFAFKCAAEELFMQLQEIOMNNSINVVEPVPVRRN--HOTELVVP 120
DB 76 SEAGRRCSGPGIYTRVHNAEQLYPMFQRYINAVNTDAFVQGERERVNSAHSVNVNG 135
QY 121 RT 122
DB 136 RT 137
RESULT 7
AAW62507
ID AAW62507 standard; protein; 481 AA.
XX
XX AAW62507;
AC
XX 16-SEP-1998 (first entry)
DT
XX p62 protein downstream of tyrosine kinases (p62-dok).
DE
XX p62-dok; protein downstream of tyrosine kinases;
KW

KW tyrosine phosphorylation; receptor tyrosine kinase; oncogenic cell;
KW p120 ras GTPase-activating protein; GAP; signal transduction pathway;
KW antibody; diagnosis; treatment; chronic myelogenous leukaemia; CML;
KW human cancer.
XX
XX Mammalia.
XX
XX WO9820129-A1.
XX
XX 14-MAY-1998.
XX
XX 30-OCT-1997; 97WO-US019788.
XX
XX 01-NOV-1996; 96US-0030418P.
XX
XX 22-JAN-1997; 97US-00787091.
XX
XX (COLD-) COLD SPRING HARBOR LAB.
XX PA (SLOK) SLOAN KETTERING INST CANCER RES.
XX
XX Carpino NA, Kobayashi R, Wisniewski DG, Strife AO, Clarkson BD;
XX
XX WPI; 1998-286945/25.
XX DR N-PSDB; AAV39415.
XX
XX Newly isolated human p62-dok protein - used to develop products for the
XX diagnosis and treatment of conditions in which oncogenic tyrosine kinase
XX is active, e.g. Chronic myelogenous leukaemia.
XX
XX Claim 1; Fig 2A-C; 86pp; English.
XX
XX The present sequence represents p62-dok (protein downstream of tyrosine
XX kinases) protein. The protein is tyrosine phosphorylated in haematopoietic
XX mammalian cells after stimulation with a ligand for a receptor tyrosine
XX kinase found in oncogenic cells. The p62-dok protein binds p120 ras
XX GTPase-activating protein (GAP). The p62-dok protein is involved in a
XX signal transduction pathway initiated by receptor tyrosine kinase, and is
XX rapidly tyrosine phosphorylated. The p62-dok protein is also present in
XX normal cells, but is not constitutively phosphorylated by normal receptor
XX tyrosine kinase. Antibodies against p62-dok and aberrantly phosphorylated
XX p62-dok, and probes derived from the p62-dok DNA sequence can be used in
XX the diagnosis and treatment of conditions in which an oncogenic tyrosine
XX kinase is active such as chronic myelogenous leukaemia (CML) and other
XX human cancers
XX
XX Sequence 481 AA;
XX
XX Query Match 20.6%; Score 143.5; DB 2; Length 481;
XX Best Local Similarity 29.5%; Pred. No. 4.4e-08;
XX Matches 33; Conservative 20; Mismatches 50; Indels 9; Gaps 2;
XX
XX 24 GSGIMELTDTELILYTRKDS-----VKWHYLCRLRYGYDSNLFSGSGRCQTGGIGF 77
XX
XX 172 GSYVLVRAERLTLLTVGAQSQILEPLLSWPTLLRYYGRDKVMFSFEAGRCPSGPTF 231
XX
XX 78 AFKCARAEELFNMLOEIMONNSINVVE---EPVVERNHHQTELEVPTPTPT 126
XX
XX 232 TPTAQGNDFQAVETAIRHQAKGQAGQGHVLRADSHGEVAEGKLPSPP 283
XX
XX
XX RESULT 8
XX ADD14099
XX ID ADD14099 standard; protein; 509 AA.
XX
XX AC ADD14099;
XX
XX 01-JAN-2004 (first entry)
XX
XX Human src biomarker polypeptide SEQ ID NO:288.
XX
XX predictor set; protein tyrosine kinase activity modulator;
XX protein tyrosine kinase pathway; protein tyrosine kinase; cytostatic;
XX gene therapy; drug sensitivity; genetic profile; cancer; human.
XX

OS Homo sapiens.
XX
XX WO2003062395-A2.
XX
XX 31-JUL-2003.
XX
XX 17-JAN-2003; 2003WO-US001981.
XX
XX 18-JAN-2002; 2002US-0350061P.
XX
XX (BRIM) BRISTOL-MYERS SQUIBB CO.
XX
XX Huang F, Fairchild CR, Lee FY, Shaw P;
XX WPI; 2003-636735/60.
XX DR N-PSDB; ADD14695.
XX
XX New polynucleotides and polypeptides for predicting the activity of
XX compounds that interact with protein tyrosine kinases and/or protein
XX tyrosine kinase pathways.
XX
XX Claim 10; SEQ ID NO 288; 139pp; English.
XX
XX The present invention describes a predictor set comprising a plurality of
XX polynucleotides or polypeptides whose expression pattern is predictive of
XX the response of cells to treatment with a compound that modulates protein
XX tyrosine kinase activity or members of the protein tyrosine kinase
XX pathway. Also described: (1) predicting whether a compound is capable of
XX modulating the activity of cells, comprising obtaining a sample of cells,
XX determining whether the cells express a plurality of markers, and
XX correlating the expression of the markers to the compound's ability to
XX modulate the activity of the cells; (2) a plurality of cell lines for
XX identifying polynucleotides and polypeptides whose expression levels
XX correlate with compound sensitivity or resistance of cells associated
XX with a disease state; and (3) identifying polynucleotides and
XX polypeptides that predict compound sensitivity or resistance of cells
XX associated with a disease state, comprising subjecting the plurality of
XX cell lines to one or more compounds, analysing the expression pattern of
XX a microarray of polynucleotides or polypeptides, and selecting
XX polynucleotides or polypeptides that predict the sensitivity or
XX resistance of cells associated with a disease state by using the
XX expression pattern of the microarray. The polynucleotides and
XX polypeptides have cytostatic activities, and can be used in gene therapy.
XX The polynucleotides and polypeptides are useful in predicting the
XX activity of compounds that interact with protein tyrosine kinases and/or
XX protein tyrosine kinase pathways. These may be used in determining drug
XX sensitivity in patients to allow the development of individualized
XX genetic profiles which aid in treating diseases and disorders (e.g.
XX cancer) based on patient response at a molecular level. The present
XX sequence is used in the exemplification of the present invention.
XX
XX Sequence 509 AA;
XX
XX Query Match 20.6%; Score 143.5; DB 7; Length 509;
XX Best Local Similarity 29.5%; Pred. No. 4.4e-08;
XX Matches 33; Conservative 20; Mismatches 50; Indels 9; Gaps 2;
XX
XX 24 GSGIMELTDTELILYTRKDS-----VKWHYLCRLRYGYDSNLFSGSGRCQTGGIGF 77
XX
XX 200 GSYVLVRAERLTLLTVGAQSQILEPLLSWPTLLRYYGRDKVMFSFEAGRCPSGPTF 259
XX
XX 78 AFKCARAEELFNMLOEIMONNSINVVE---EPVVERNHHQTELEVPTPTPT 126
XX
XX 260 TPTAQGNDFQAVETAIRHQAKGQAGQGHVLRADSHGEVAEGKLPSPP 311
XX
XX
XX RESULT 9
XX ABG66871
XX ID ABG66871 standard; protein; 541 AA.
XX
XX AC ABG66871;
XX
XX 24-SEP-2002 (first entry)
XX

XX DE Human prostate specific protein DEX0283_179.
 XX KW Human; prostate specific protein; cytostatic; PSP; prostate cancer;
 XX KW Gene therapy; vaccine; non-cancerous prostatic disease.
 XX OS Homo sapiens.
 XX FN WO200242329-A2.
 XX PD 30-MAY-2002.
 XX XX
 XX PF 20-NOV-2001; 2001WO-US045177.
 XX XX
 XX PR 21-NOV-2000; 2000US-0252189P.
 XX XX
 XX PA (DIAD-) DIADEXUS INC.
 XX XX
 XX PI Salceda S, Macina RA, Recipon H, Sun Y, Liu C;
 XX DR WPI; 2002-500278/53.
 XX XX
 XX PT New prostate specific polypeptides useful for identifying, diagnosing,
 XX PT monitoring, staging, imaging and treating prostate cancer and non-
 XX PT cancerous disease states in prostate.
 XX PS Claim 11; Page 230-233; 248pp; English.
 XX CC
 XX CC The invention relates to an isolated polypeptide that comprises a
 XX CC sequence with 60 % sequence identity to one of 103 prostate specific
 XX CC polypeptide (PSP) sequences (S1), given in specification, or a sequence
 XX CC encoded by a nucleic acid comprising one of 114 prostate specific nucleic
 XX CC acid (PSNA) sequences (S2), given in specification. Also included are a
 XX CC vector comprising the PSNA, a host cell comprising the vector,
 XX CC preparation of the PSP, a polypeptide encoded by the PSNA; an anti-PSP
 XX CC antibody the antibody or its fragment, a kit for detecting a risk of
 XX CC cancer or presence of cancer in a patient, comprising a unit for
 XX CC determining the presence of the PSP or the PSNA in a sample of a patient
 XX CC and a vaccine comprising the PSP or the PSNA. The PSNA is useful for
 XX CC determining the presence of a prostate specific nucleic acid (PSNA) in a
 XX CC sample. The antibody to the PSP is useful for determining the presence of
 XX CC a prostate specific protein in a sample. The antibody is also useful for
 XX CC treating a patient with prostate cancer, where the administration of the
 XX CC antibody induces an immune response against the prostate cancer cell
 XX CC expressing the nucleic acid molecule or polypeptide. The PSP or the PSNA
 XX CC is useful for diagnosing and monitoring the presence and metastases of
 XX CC prostate cancer in a patient. The PSP is useful as a vaccine component
 XX CC for eliciting a humoral and/or cellular immune response, and for
 XX CC producing transgenic animals and cells, and for producing engineered
 XX CC prostate tissue for treatment and research. The PSNA is useful for
 XX CC detecting prostate cancer by detecting genetic lesions or mutations. The
 XX CC PSP and the PSNA are useful for detecting, diagnosing, monitoring
 XX CC staging and predicting cancers. The PSP and the PSNA are also useful for
 XX CC identifying prostate tissue, and thus is useful in forensic science,
 XX CC prostate cell differentiation and development and in tissue engineering.
 XX CC The PSP, the PSNA and the antibody are useful for identifying,
 XX CC diagnosing, monitoring, staging, imaging and treating prostate cancer and
 XX CC non-cancerous disease states in prostate tissue. The present sequence is
 XX CC a prostate specific protein (PSP) sequence
 XX SQ Sequence 541 AA;

Query Match 20.6%; Score 143.5; DB 5; Length 541;
 Best Local Similarity 29.5%; Pred. No. 4.8e-08;
 Matches 33; Conservative 20; Mismatches 50; Indels 9; Gaps 2;

QY 24 GSGIMELTDTLLVTRKDS-----VKWYHLCIRYGYDSNLFPSGRCRCOTGGGIF 77
 DB 248 GSYLVRAEALITLLTVGAQSQILEPLLSWYTLRLRYGRDKVMFSEAGRCRCSGGPTF 307
 QY 78 AFKCAAEELNMLQEIQNNSINVE---EPVYERNNHQTELEVPRTPR 126

Db 308 TFOAQNDIFQAVETAHROKQKQAGHDVLRADSHGEVAEGKLPSP 359

RESULT 10
 AAY79184
 ID AAY79184 standard; protein; 359 AA.
 XX AC AAY79184;
 XX DT 19-JUN-2000 (first entry)
 XX DE Haematopoietic stem cell specific protein.
 XX KW Haematopoietic stem cell; immune system disorder; leukaemia;
 XX KW antileukaemic; immunomodulator; therapy; mouse.
 XX OS Mus musculus.
 XX FH Key Location/Qualifiers
 XX FT Misc-difference 344
 XX FT /note= "nucleotide sequence encoding this protein has an
 XX FT in-frame stop codon following Leu codon 343"
 XX PN WO200011168-A2.
 XX XX
 XX PD 02-MAR-2000.
 XX PF 20-AUG-1999; 99WO-US019052.
 XX PR 21-AUG-1998; 98US-00138132.
 XX PA (UYPR-) UNIV PRINCETON.
 XX PI Lemischka I, Moore K;
 XX DR WPI; 2000-237650/20.
 XX DR N-PSDB; AA294125.
 XX PT Hematopoietic stem cell signaling proteins modulating replication and
 XX PT differentiation for treating immune system disorders and leukemia.
 XX PS Claim 21; Page 232-233; 256pp; English.

The present sequence is that of a mouse haematopoietic stem cell (HSC) specific protein. It is an example of claimed HSC-specific proteins (see AAY79175-93) predicted from novel isolated HSC-specific nucleic acids (see AA294077-131). The HSCs are especially primitive HSCs (PHSCs) such as umbilical cord cells, bone marrow cells and foetal liver cells. The encoded proteins are growth factors, transcription factors, splicing factors, capping factors, transport proteins, translation factors or replication factors that modulate HSC activity, especially for identifying PHSC-specific nucleic acids; for generating a stem cell/progenitor cell from PHSCs; for identifying the presence of a PHSC in a sample; for identifying the presence in a sample of a compound that modulates HSC activity; for using such a compound to treat an immune system condition, especially leukaemia; for introducing exogenous nucleic acid into a HSC; and for ex vivo expansion of HSCs. Also claimed are vectors, host cells, and an antibody that specifically binds a an HSC-specific protein

XX SQ Sequence 359 AA;

Query Match 20.3%; Score 141; DB 3; Length 359;
 Best Local Similarity 28.3%; Pred. No. 5.5e-08;
 Matches 36; Conservative 20; Mismatches 47; Indels 24; Gaps 3;

QY 23 LGSIMELTDTLLVTRKDSVKWYHLCIRYGYDSNLFPSGRCRCOTGGGIFAFKCA 82
 DB 184 LGQDDIQLRET-----SKPQACFSWFYFLRYKYGDKGVFEAGRCDSGEGLFAPSP 238
 QY 83 RABEELPNMLQEIQNNSINVE-----EPVYERNNHQTELEVPRTPR 124

Db 239 RAPDICGVAAIARQERLPELAMPPCPLRALSLPSLEPPGELREVAPGFELTPRK 298
 QY 125 TP-TTPG 130
 Db 299 LPLTDPG 305
 ABR83336
 ID ABR83336 standard; protein; 331 AA.
 AC ABR83336;
 XX
 DT 06-OCT-2003 (first entry)
 DE Human NOV3a protein SEQ ID NO:6.
 XX
 KW Human; NOVX; cytostatic; antidiabetic; neuroprotective; antiparkinsonian;
 KW anorectic; gene therapy; vaccine; cancer; neurodegenerative disorder;
 KW Parkinson's disease; metabolic disorder; diabetes; obesity;
 KW tissue typing.
 XX
 OS Homo sapiens.
 XX
 PN WO2003052061-A2.
 XX
 PD 26-JUN-2003.
 XX
 PF 03-DEC-2002; 2002WO-US038821.
 XX
 PR 17-DEC-2001; 2001US-0341477P.
 PR 17-DEC-2001; 2001US-0341540P.
 PR 20-DEC-2001; 2001US-0342592P.
 PR 31-DEC-2001; 2001US-0344903P.
 PR 17-APR-2002; 2002US-0373288P.
 PR 15-MAY-2002; 2002US-0380981P.
 PR 17-MAY-2002; 2002US-0381495P.
 PR 28-MAY-2002; 2002US-0383744P.
 PR 23-MAY-2002; 2002US-0384024P.
 PR 07-AUG-2002; 2002US-0401788P.
 PR 26-AUG-2002; 2002US-0406353P.
 PR 31-OCT-2002; 2002US-0422756P.
 PR 03-DEC-2002; 2002US-00307928.
 XX
 PA (CURA-) CURAGEN CORP.
 XX
 PI Alsobrook JP, Anderson DW, Boldog FL, Burgess CE, Catterton E;
 PI Edinger SR, Gorman L, Guo X, Ji W, Kekuda R, Li L, Patraajan M;
 PI Rieger DK, Shenoy SG, Seytek KA, Vernet CAV, Voss EZ, Zhong M;
 XX
 DR WPI: 2003-533005/50.
 DR N-PSDB; ACF06235.
 XX
 PT New NOVX polypeptide, useful for preparing a composition for treating or
 PT preventing e.g. cancer, neurodegenerative disorders such as Parkinson's
 PT disease, or metabolic disorders such as diabetes or obesity, or for
 PT tissue typing.
 XX
 PS Claim 1; Page 102; 190pp; English.
 XX
 CC ACF06233 to ACF06242 encode the human NOVX proteins given in ABR83334 to
 CC ABR83343, designated NOV1a, NOV2a, NOV3a, NOV4a, NOV4b, NOV5a, NOV6a,
 CC NOV7a, NOV8a and NOV9a respectively. NOVX sequences can have cytostatic,
 CC antidiabetic, neuroprotective, antiparkinsonian and anorectic activities,
 CC and can be used in vaccines and gene therapy. The NOVX polypeptides can
 CC be used for preparing a composition for treating or preventing a
 CC pathology associated with the NOVX-polypeptides e.g. cancer,
 CC neurodegenerative disorders such as Parkinson's disease, or metabolic
 CC disorders such as diabetes or obesity, or for tissue typing
 XX
 SQ Sequence 331 AA;
 Query Match 19.8%; Score 137.5; DB 6; Length 331;

Best Local Similarity 31.5%; Pred. No. 1.3e-07;
 Matches 34; Conservative 19; Mismatches 42; Indels 13; Gaps 3;
 QY 24 GSGIMELTDTELILYTRKRSVK---WHYLCRRYGYDSNLFESGRCOTGQIFAFK 80
 Db 153 GECTMQITHENILYLDIHNAKVKLVMPPLSSLRRYGRDSTWFTESGRMCDTGGLFTFQ 212
 QY 81 CARAELEFNMQEIMONNSINVEE-----PVVERNNHQTELEVPR 122
 Db 213 TREGE---MTYQKVHSATLAIAGQHERLMLEMEQKARLOTSLTEPMT 256
 RESULT 12
 AAM39806
 ID AAM39806 standard; protein; 337 AA.
 AC AAM39806;
 XX
 DT 22-OCT-2001 (first entry)
 DE Human polypeptide SEQ ID NO 2951.
 XX
 KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
 KW peripheral nervous system; neuropathy; central nervous system; CNS;
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
 KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
 KW leukaemia.
 XX
 OS Homo sapiens.
 XX
 PN WO200153312-A1.
 XX
 PD 26-JUL-2001.
 XX
 PF 26-DEC-2000; 2000WO-US034263.
 XX
 PR 23-DEC-1999; 99US-00471275.
 PR 21-JAN-2000; 2000US-00488725.
 PR 25-APR-2000; 2000US-00552317.
 PR 20-JUN-2000; 2000US-00598042.
 PR 19-JUL-2000; 2000US-00820312.
 PR 03-AUG-2000; 2000US-00853450.
 PR 14-SEP-2000; 2000US-00662191.
 PR 19-OCT-2000; 2000US-00693036.
 PR 29-NOV-2000; 2000US-00727344.
 XX
 XX (HYSE-) HYSEQ INC.
 XX
 PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
 PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;
 PI Zhou P, Goodrich R, Drmanac RT;
 XX
 DR WPI: 2001-442253/47.
 DR N-PSDB; AAI58962.
 XX
 PT Novel nucleic acids and polypeptides, useful for treating disorders such
 PT as central nervous system injuries.
 XX
 PS Example 4; SEQ ID NO 2951; 10078pp; English.
 XX
 CC The invention relates to human nucleic acids (AAI57798-AAI61369) and the
 CC encoded polypeptides (AAM39642-AAM42213) with nootropic,
 CC immunosuppressant and cytostatic activity. The polynucleotides are useful
 CC in gene therapy. A composition containing a polypeptide or polynucleotide
 CC of the invention may be used to treat diseases of the peripheral nervous
 CC system, such as peripheral nervous injuries, peripheral neuropathy and
 CC localised neuropathies and central nervous system diseases, such as
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
 CC utilisation of the activities such as: immune system suppression,
 CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,

CC assays for receptor activity, arthritis and inflammation, leukaemias and
 CC C.N.S disorders. Note: The sequence data for this patent did not form
 CC part of the printed specification

XX SQ Sequence 337 AA;

Query Match 19.8%; Score 137.5; DB 4; Length 337;
 Best Local Similarity 31.5%; Pred. No. 1.3e-07;
 Matches 34; Conservative 19; Mismatches 42; Indels 13; Gaps 3;

OY 24 GSGIMELTDTLLIYTRKDSVK---WHYLCRLRYGYDSNLFSPESGRRCQTGQIFAFK 80
 DB 153 GECTMQITHENIYLDIHNAKVLVMPPLSLRLRYGRDSTWTFESGRMCDTGESLFTFQ 212

OY 81 CARAEELFNLMQEIOMNNSINWVEE-----PVVERNNHOTELEVPRT 122

DB 213 TREGE---MIYQKVHSATLAIHQERHLEMLEMEQKARLQTSLEPMT 256

RESULT 13

AAM41592

ID AAM41592 standard; protein; 365 AA.

XX AC AAM41592;

XX 22-OCT-2001 (first entry)

XX Human polypeptide SEQ ID NO 6523.

XX Human; nontropic; immunosuppressant; cytostatic; gene therapy; cancer;
 KW peripheral nervous system; neuropathy; central nervous system; CNS;
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
 KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
 KW leukaemia.

XX OS Homo sapiens.

XX WO200153312-A1.

XX 26-JUL-2001.

XX 26-DEC-2000; 2000WO-US034263.

XX 23-DEC-1999; 99US-00471275.

XX 21-JAN-2000; 2000US-00488725.

XX 25-APR-2000; 2000US-00552317.

XX 20-JUN-2000; 2000US-00598042.

XX 19-JUL-2000; 2000US-00620312.

XX 03-AUG-2000; 2000US-00653450.

XX 14-SEP-2000; 2000US-00662191.

XX 19-OCT-2000; 2000US-00693036.

XX 29-NOV-2000; 2000US-00727344.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
 PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;
 PI Zhou P, Goodrich R, Drmanac RT;

XX WPI; 2001-442253/47.

XX N-PSDB; AAI60748.

XX Novel nucleic acids and polypeptides, useful for treating disorders such
 PT as central nervous system injuries.

XX Example 2; SEQ ID NO 6523; 10078pp; English.

XX The invention relates to human nucleic acids (AAI57798-AAI61369) and the
 CC encoded polypeptides (AAM38642-AAI42213) with nontropic,
 CC immunosuppressant and cytostatic activity. The polynucleotides are useful
 CC in gene therapy. A composition containing a polypeptide or polynucleotide
 CC of the invention may be used to treat diseases of the peripheral nervous

CC system, such as peripheral nervous injuries, peripheral neuropathy and
 CC localised neuropathies and central nervous system diseases, such as
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
 CC utilisation of the activities such as: Immune system suppression,
 CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
 CC assays for receptor activity, arthritis and inflammation, leukaemias and
 CC C.N.S disorders. Note: The sequence data for this patent did not form
 CC part of the printed specification

XX SQ Sequence 365 AA;

Query Match 19.8%; Score 137.5; DB 4; Length 365;

Best Local Similarity 31.5%; Pred. No. 1.5e-07;

Matches 34; Conservative 19; Mismatches 42; Indels 13; Gaps 3;

OY 24 GSGIMELTDTLLIYTRKDSVK---WHYLCRLRYGYDSNLFSPESGRRCQTGQIFAFK 80
 DB 181 GECTMQITHENIYLDIHNAKVLVMPPLSLRLRYGRDSTWTFESGRMCDTGESLFTFQ 240

OY 81 CARAEELFNLMQEIOMNNSINWVEE-----PVVERNNHOTELEVPRT 122

DB 241 TREGE---MIYQKVHSATLAIHQERHLEMLEMEQKARLQTSLEPMT 284

RESULT 14

AAM41593

ID AAM41593 standard; protein; 365 AA.

XX AC AAM41593;

XX 22-OCT-2001 (first entry)

XX Human polypeptide SEQ ID NO 6524.

XX Human; nontropic; immunosuppressant; cytostatic; gene therapy; cancer;
 KW peripheral nervous system; neuropathy; central nervous system; CNS;
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
 KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
 KW leukaemia.

XX OS Homo sapiens.

XX WO200153312-A1.

XX 26-JUL-2001.

XX 26-DEC-2000; 2000WO-US034263.

XX 23-DEC-1999; 99US-00471275.

XX 21-JAN-2000; 2000US-00488725.

XX 25-APR-2000; 2000US-00552317.

XX 20-JUN-2000; 2000US-00598042.

XX 19-JUL-2000; 2000US-00620312.

XX 03-AUG-2000; 2000US-00653450.

XX 14-SEP-2000; 2000US-00662191.

XX 19-OCT-2000; 2000US-00693036.

XX 29-NOV-2000; 2000US-00727344.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
 PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;
 PI Zhou P, Goodrich R, Drmanac RT;

XX WPI; 2001-442253/47.

XX N-PSDB; AAI60749.

XX Novel nucleic acids and polypeptides, useful for treating disorders such
 PT as central nervous system injuries.

PS Example 2; SEQ ID NO 6524; 10078pp; English.

XX The invention relates to human nucleic acids (AAI57798-AAI61369) and the

CC encoded polypeptides (AAM38642-AAM42213) with neurotropic,

CC immunosuppressant and cytostatic activity. The polynucleotides are useful

CC in gene therapy. A composition containing a polypeptide or polynucleotide

CC of the invention may be used to treat diseases of the peripheral nervous

CC system, such as peripheral nervous injuries, peripheral neuropathy and

CC localised neuropathies and central nervous system diseases, such as

CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic

CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the

CC utilisation of the activities such as: Immune system suppression,

CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic

CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,

CC assays for receptor activity, arthritis and inflammation, leukaemia and

CC C.N.S disorders. Note: The sequence data for this patent did not form

CC part of the printed specification

XX

SQ Sequence 365 AA;

Query Match 19.8%; Score 137.5; DB 4; Length 365;

Best Local Similarity 31.5%; Pred. No. 1.5e-07;

Matches 34; Conservative 19; Mismatches 42; Indels 13; Gaps 3;

QY 24 GSGIMELDTLILYTRKRSVK---WHYLCRLRYGYDSNLFESGRCRCOTGQGFAPK 80

DB 181 GECTMQITHENIYLDIHNAKVLVWVPLSLRLRYGRDSTWTFESGRMCDTGEGLFTFQ 240

QY 81 CARAEELFNLMQIMQNSINVVE-----FVVERNHHOTELEVPT 122

DB 241 TRGE-----MIYQKVSHTLAIAEQHERLMEQKARLQTSLEPT 284

RESULT 15

AAAB52062

ID AAB52062 standard; protein; 412 AA.

XX AC AAB52062;

XX 21-FEB-2001 (first entry)

XX Gene 2 human secreted protein homologous amino acid sequence #111.

XX Human; secreted protein; cytostatic; immunosuppressive; neurotropic;

XX neuroprotective; antiviral; antiallergic; hepatotropic; antidiabetic;

XX antiinflammatory; antitumor; vulnerable; anticonvulsant; antibacterial;

XX antifungal; antiparasitic; cardiac; cancer; immune disease; allergy;

XX cardiovascular disorder; wound healing; infection; neurological disease.

XX Homo sapiens.

XX WO200061596-A1.

XX 19-OCT-2000.

XX 06-APR-2000; 2000WO-US008983.

XX 09-APR-1999; 99US-0128703P.

XX 14-JAN-2000; 2000US-0176068P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX (ROSE/) ROSEN C A.

XX Ruben SM, Komatsoulis G;

XX WPI; 2000-611865/59.

XX Fifty nucleic acid molecules encoding human secreted proteins, useful in

XX the prevention, treatment and diagnosis of cancer, immune disorders,

XX cardiovascular disorders and neurological diseases.

XX Disclosure; Page 11; 505pp; English.

XX

CC Polynucleotide sequences AAC95521 - AAC95570 represent cDNA encoding 50

CC human secreted proteins AAB52012 - AAB52161. Sequences AAB52062 -

CC AAB52103 represent alternative polypeptides encoded by the genes, and

CC amino acid sequences with which they share homology. The genes and

CC proteins have activities dependent on the tissues and cells in which they

CC are expressed. Examples of their activities include cytostatic;

CC immunosuppressive; neurotropic; neuroprotective; antiviral; antiallergic;

CC hepatotropic; antidiabetic; antiinflammatory; antitumor; vulnerable;

CC anticonvulsant; antibacterial; antifungal; antiparasitic; and cardiac.

CC The secreted proteins, polypeptides, antagonists and agonists may be

CC useful in treating, preventing and/or diagnosing diseases and disorders

CC such as cancer, particularly breast and ovarian cancer, and other cancers

CC of the adrenal gland, bone, bone marrow, breast, gastrointestinal tract,

CC liver, lung, or urogenital. Immune disorders such as Addison's disease,

CC allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis,

CC diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid

CC arthritis and ulcerative colitis; cardiovascular disorders such as

CC myocardial ischaemias; wound healing; neurological diseases such as

CC cerebral anoxia and epilepsy; and infectious diseases such as viral,

CC bacterial, fungal and parasitic infections may also be treated using the

CC proteins and polynucleotides of the invention. Sequences AAC95512 -

CC AAC95520 and AAB52011 are used in the isolation and characterisation of

CC the proteins and polynucleotides of the invention

XX

SQ Sequence 412 AA;

Query Match 19.4%; Score 135; DB 3; Length 412;

Best Local Similarity 34.0%; Pred. No. 3.5e-07;

Matches 33; Conservative 13; Mismatches 33; Indels 18; Gaps 3;

QY 47 WHYLCRLRYGYDSNLFESGRCRCOTGQGFAPKCARAEELFNLMQIM--QNSINVVE 104

DB 194 WYRFLRRFGDKVTFEAGRCRCVSGEGNEFETRCQNEIFLAELEAISAKQKAPATP 253

QY 105 FVVERNHHOTELEVPR-----TPRTPTTP 129

DB 254 QP-----QPATIPASLPRDPSYSRPHDSLPPSPPTP 286

Search completed: August 2, 2004, 09:29:30

Job time : 84.1053 secs

Result No.	Score	Query		Length	DB	ID	Description
		Match	%				
1	178	25.6	195	2	T32819		hypothetical prote
2	178	17.9	265	2	T08731		hypothetical prote
3	82	11.8	1235	1	S16948		insulin receptor s
4	78	11.2	1231	1	S30185		insulin receptor s
5	78	11.2	1242	2	JS0670		insulin receptor s
6	75.5	10.8	320	2	T48191		hypothetical prote
7	75.5	10.8	630	2	T03052		hypothetical prote
8	75.5	10.8	753	2	C96668		unknown protein F1
9	75	10.8	435	2	C39171		cyclin A/CDK3-asso
10	73	10.5	893	2	A47550		bride of sevenless
11	72.5	10.4	1528	2	T37308		Apfase homolog - C
12	72	10.3	189	2	T04499		auxin-induced prot
13	72	10.3	314	2	D70313		riboflavin kinase
14	72	10.3	359	2	T37958		probable myotubula
15	72	10.3	1310	2	T30562		resistance protein
16	71.5	10.3	1309	1	BV8ID9		RAD9 protein - yea
17	71	10.2	244	2	D71608		hypothetical prote
18	70.5	10.1	351	2	H72360		divalent cation tr
19	70	10.1	275	2	S77476		tRNA-pseudouridine
20	70	10.1	522	2	C90073		hypothetical prote
21	70	10.1	1240	2	JCS209		insulin receptor s
22	69.5	10.0	512	2	R86196		hypothetical prote
23	69.5	10.0	612	1	W1WLE		E1 protein - human
24	69.5	10.0	924	2	T05626		hypothetical prote
25	69.5	10.0	2823	2	T23064		hypothetical prote
26	69.5	10.0	2823	2	F87908		protein T22A3.8 [1
27	69.5	10.0	3102	2	T43291		laminin alpha chai
28	69	9.9	161	2	C84769		hypothetical prote
29	69	9.9	113	2	G64606		hydantoin utilizat

Best Local Similarity 26.3%; Pred. No. 1.8e-05;
Matches 30; Conservative 29; Mismatches 40; Indels 15; Gaps 3;
QY 24 GSGIMEJTDTELLYTRKRSVK---WHYLCLRRYGYDSNLFSPESGRRCQTGGIFAFK 80
D 112 GECALQITYEYICLDVNDQVRKLSWPLSALGRYGRDITFTFEAGRMCETGEGLIFFQ 171
QY 81 CARAEELFNML-----QEIWNNSINNVPEPVVER-NNHQTELEVPR 122
D 172 TRGEAIYQVHSAALIAISOHERLLQSVKNSMLQMKMSERAAASLSTWFLPRS 225
RESULT 3
insulin receptor substrate IRS-1 - rat
N:Alternate names: insulin receptor substrate pp185
C:Species: Rattus norvegicus (Norway rat)
C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 18-Jun-1999
C:Accession: S16948; A39811
R:Sun, X.J.; Rothenberg, P.; Kahn, C.R.; Backer, J.M.; Azaki, E.; Wilden, P.A.; Cahill,
Nature 352, 73-77, 1991
A:Title: Structure of the insulin receptor substrate IRS-1 defines a unique signal trans
A:Reference number: S16948; MUID:91287824; PMID:1648180
A:Accession: S16948
A:Molecule type: mRNA
A:Residues: 1-1235 <SUN>
A:Cross-references: EMBL:X58375; NID:956503; PIDN:CAA41264.1; PID:956504
R:Rothenberg, P.L.; Lane, W.S.; Karaik, A.; Backer, J.; White, M.; Kahn, C.R.
J. Biol. Chem. 266, 8302-8311, 1991
A:Title: Purification and partial sequence analysis of pp185, the major cellular substra
A:Reference number: A39811; MUID:91217066; PMID:2022647
A:Accession: A39811
A:Molecule type: protein
A:Residues: 44-51,173-178;203-243;489-506;635-646;871-882,'I',884,'X',886-888;932-936,'X'
A:Note: The phosphotyrosine residue was not identified
C:Comment: This protein and the beta chain of the insulin receptor itself are the major
C:Comment: Phosphorylation of this protein in response to insulin is maximal at 30 second
insulin.
C:Superfamily: insulin receptor substrate IRS-1; pleckstrin repeat homology
C:Keywords: phosphoprotein; signal transduction
F:11-113/Domain: pleckstrin repeat homology <PLK>
F:872-891/Region: glutamine-rich
Query Match 11.8%; Score 82; DB 1; Length 1235;
Best Local Similarity 30.0%; Pred. No. 3.8;
Matches 24; Conservative 7; Mismatches 21; Indels 28; Gaps 2;
QY 52 LRRYGYDSNLFSPESGRRCQTGGIFAFKCARAEELFNMLQEIWMQNSINNVPEPVVER 111
D 206 IRRCGHSENFIEVGRSAVTGPFQFNWQ-----VDSVVAQN 243
QY 112 NHQTELEVPR-----PRT 125
D 244 MHEITLQAMRAMSDEPRPT 263
RESULT 4
insulin receptor substrate 1 - mouse
C:Species: Mus musculus (house mouse)
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 16-Jul-1999
C:Accession: S30185; S43514
R:Keller, S.R.; Aberkand, R.; Garner, C.W.; Lienhard, G.E.
Biochim. Biophys. Acta 1172, 323-326, 1993
A:Title: The insulin-elicited 160 kDa phosphotyrosine protein in mouse adipocytes is an
A:Reference number: S30185; MUID:93192326; PMID:8448209
A:Accession: S30185
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1231 <REL>
A:Cross-references: EMBL:X69722; NID:9297913; PIDN:CAA49378.1; PID:9297914
R:Araki, E.; Haag III, B.L.; Kahn, C.R.
Biochim. Biophys. Acta 1221, 353-356, 1994

A:Title: Cloning of the mouse insulin receptor substrate-1 (IRS-1) gene and complete seq.
A:Reference number: S43514; MUID:94220494; PMID:8167159
A:Accession: S43514
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1037,'AS',1038-1179,'H',1181-1231 <ARA>
A:Cross-references: EMBL:L24563; NID:9407993; PIDN:AAA39335.1; PID:9407994
C:Superfamily: insulin receptor substrate IRS-1; pleckstrin repeat homology
F:11-113/Domain: pleckstrin repeat homology <PLK>
Query Match 11.2%; Score 78; DB 2; Length 1231;
Best Local Similarity 30.0%; Pred. No. 10;
Matches 21; Conservative 7; Mismatches 20; Indels 22; Gaps 1;
QY 52 LRRYGYDSNLFSPESGRRCQTGGIFAFKCARAEELFNMLQEIWMQNSINNVPEPVVER 111
D 206 IRRCGHSENFIEVGRSAVTGPFQFNWQ-----VDSVVAQN 243
QY 112 NHQTELEVPR 121
D 244 MHEITLQAMR 253
RESULT 5
insulin receptor substrate-1 - human
C:Species: Homo sapiens (man)
C:Date: 30-Jun-1992 #sequence_revision 23-Aug-1996 #text_change 16-Jul-1999
C:Accession: I53160; J50670; PNO678
R:Araki, E.; Sun, X.J.; Haag, B.L.; Chuang, L.; Zhang, Y.; Yang-Feng, T.L.; White, M.; Ka
Diabetes 42, 1041-1054, 1993
A:Title: Human skeletal muscle insulin receptor substrate-1. Characterization of the cdw
A:Reference number: I53160; MUID:93292738; PMID:8513971
A:Accession: I53160
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1242 <RES>
A:Cross-references: GB:S25339; NID:9386256; PIDN:AA27175.1; PID:9386257
R:Nishiyama, M.; Wands, J.R.
Biochem. Biophys. Res. Commun. 183, 280-285, 1992
A:Title: Cloning and increased expression of an insulin receptor substrate-1-like gene in
A:Reference number: J50670; MUID:92181456; PMID:1311924
A:Accession: J50670
A:Molecule type: mRNA
A:Residues: 1-134, 'G', 135-361, 'R', 363-383, 'R', 385-1242 <NIS>
A:Cross-references: GB:S85963; NID:9246485; PIDN:AA521608.1; PID:9246486
A:Experimental source: hepatocellular carcinoma cell line FOCUS
R:Smith, L.K.; Bradshaw, M.; Croall, D.E.; Garner, C.W.
Biochem. Biophys. Res. Commun. 196, 767-772, 1993
A:Title: The insulin receptor substrate (IRS-1) is a PEST protein that is susceptible to
A:Reference number: PNO678; MUID:94059102; PMID:8240352
A:Accession: PNO678
A:Molecule type: protein
A:Residues: 1-14,338-353;411-439;538-545,'V',547-567;656-697;724-758;932-943;1028-1056;11
C:Comment: This protein appears to be the major substrate for insulin-stimulated tyrosine
o dock various proteins containing the phosphotyrosine-binding Src-homology domain 2 (SH;
C:Comment: This protein contains at least 11 PEST regions, which suggests rapid turnover.
C:Genetics:
A:Gene: GDB:IRS1
A:Cross-references: GDB:I33974; OMIM:147545
A:Map position: 2q36-2q36
A:Introns: #status absent
A:Superfamily: insulin receptor substrate IRS-1; pleckstrin repeat homology
C:Keywords: ATP; phosphoprotein
F:11-113/Domain: pleckstrin repeat homology <PLK>
F:46,465,555,612,632,662,732,941,989,1012/Binding site: phosphate (Tyr) (covalent) #statu
F:78,527,1100,1223/Binding site: phosphate (Ser) (covalent) (by CAMP-dependent kinase) #s
F:161/Binding site: ATP (lys) #status predicted
F:189,323,441,624,636,795,920,984,1084,1218/Binding site: phosphate (Ser) (covalent) (by
F:300,351,774/Binding site: phosphate (Thr) (covalent) (by protein kinase C) #status pred
Query Match 11.2%; Score 78; DB 2; Length 1242;
Best Local Similarity 30.0%; Pred. No. 10;

Matches 21; Conservative 7; Mismatches 20; Indels 22; Gaps 1;
QY 52 LRRYDYNLSPFSGRRCTGQGFAPKCARAELENMLOEIMONNSINVVPEPVN 111
Db 211 IRRCHSENFFIEVGRSAVTGPGFWMQ-----VDSVVAQN 248
QY 112 NHOETEVR 121
Db 249 MHTILEAMR 258

RESULT 6
T48191
hypothetical protein F7A7.220 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 28-Jul-2000
C:Accession: T48191
R:Bevan, M.; Terry, N.; Ardiles, W.; Buysshaert, C.; Dasseville, R.; De Clerck, R.; De
ewes, H.W.; Rudd, S.; Lemcke, K.; Mayer, K.F.X.
A:Reference number: 224487
A:Submitted to the Protein Sequence Database, March 2000
A:Accession: T48191
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-320 <BEV>
A:Cross-references: EMBL:AL161946
A:Experimental source: cultivar Columbia; BAC clone F7A7
C:Genetics:
A:Map position: 5
A:Introns: 68/3; 95/3; 138/3; 177/1; 251/3; 315/3
A:Note: F7A7.220
C:Superfamily: Arabidopsis thaliana hypothetical protein F7A7.220

Query Match 10.8%; Score 75.5; DB 2; Length 320;
Best Local Similarity 27.1%; Pred. No. 3.4;
Matches 26; Conservative 11; Mismatches 34; Indels 25; Gaps 3;
QY 18 DGNELGSGIMELTDIELLYTRKDSVKWHYLCRR-----YGYDNL 61
Db 18 DGNELGSGIMELTDIELLYTRKDSVKWHYLCRR-----YGYDNL 61
QY 62 FPFSGRRCT--GQGFAPKCARAELENMLOEIM 95
Db 78 F-----CGVFDGHPMGKISHHELFREFEDIL 106

RESULT 7
T00352
hypothetical protein KIAA0701 - human (fragment)
C:Species: Homo sapiens (man)
C>Date: 01-Feb-1999 #sequence_revision 01-Feb-1999 #text_change 21-Jul-2000
C:Accession: T00352
R:Ishikawa, K.; Nagase, T.; Suyama, M.; Miyajima, N.; Tanaka, A.; Kotani, H.; Nomura, N.
DNA Res. 5, 169-176, 1998
A:Title: Prediction of the coding sequences of unidentified human genes. X. The complete
A:Reference number: Z14142; MUID:98403880; PMID:9734811
A:Accession: T00352
A:Status: Preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-630 <ISH>
A:Cross-references: EMBL:AB014601; NID:G3327215; PIDN:BAA31676.1; PID:G3327216
A:Experimental source: brain; clone HG01611
C:Genetics:
A:Note: KIAA0701

Query Match 10.8%; Score 75.5; DB 2; Length 630;
Best Local Similarity 25.0%; Pred. No. 8;
Matches 36; Conservative 19; Mismatches 54; Indels 35; Gaps 6;
QY 1 DTVPDNRHNRKFKVINVDGNEI-----GSGIM--ELTDIELLYTRKDSVKWHYLCRR 54
Db 131 DYLSDKHLGKIS-----EDSSGLVYKSGGIEGISTD-----KDSFTYDSSSILN 178

QY 55 YGYDYNLSPFES-----GRRCTGQGFAPKCARAELENMLOEIMONNSIN 102
Db 179 YREDSNLSFDGNGNLSLTLSKGNETIESIF-----KAEDLLPEAASLENLISK 233
QY 103 VEEPVVNRNHOETEVRTPRT 126
Db 234 EETPPVTRLKSSQSSLSGKPKRCP 257

RESULT 8
C96668
unknown protein F15H21.18 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C:Accession: C96668
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
ansen, N.F.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, T.H.; Dewar, K.;
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.;
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, Z.A.; Lueros, J.S.; Mafti, R.; Marziani,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, I.
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: C96668
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-753 <STO>
A:Cross-references: GB:AE005173; NID:G10645404; PIDN:AAG21522.1; GSPDB:GN00141
C:Genetics:
A:Gene: F15H21.18
A:Map position: 1

Query Match 10.8%; Score 75.5; DB 2; Length 753;
Best Local Similarity 24.4%; Pred. No. 10;
Matches 33; Conservative 26; Mismatches 39; Indels 37; Gaps 7;
QY 6 NHRNK-----FKVINVDGNGELGSGIMELTDIELLYTRKDSVKWHYLCRRYGYDS 59
Db 31 NHYTKSGMSYEMMLADEGNGIKQGIK-----EHLKLQSDDELAKHFNIRAY---N 83
QY 60 NLFSPES-----GRRCTGQGFAPK-----CARAELENMLOEIMONNS 99
Db 84 MLFSTISGKVDHCLPKGR-----GPNMFAIQGALKPKSVAKAFQQLYIVDTENEVNR 139
QY 100 INVVEPVVNRNHO 114
Db 140 YNIMELRYIKNNQPQ 154

RESULT 9
I39171
cyclin A/CDK2-associated p45 - human
C:Species: Homo sapiens (man)
C>Date: 23-Feb-1996 #sequence_revision 23-Feb-1996 #text_change 21-Jul-2000
C:Accession: I39171
R:Zhang, H.; Kobayashi, R.; Galaktionov, K.; Beach, D.
Cell 82, 915-925, 1995
A:Title: p19Skp1 and p45Skp2 are essential elements of the cyclin A-CDK2 S phase kinase.
A:Reference number: I39170; MUID:96016087; PMID:753852
A:Accession: I39171
A:Status: Preliminary
A:Molecule type: mRNA
A:Residues: 1-435 <RES>
A:Cross-references: EMBL:U33761; NID:G995825; PIDN:AAC50242.1; PID:G995826
C:Genetics:
A:Gene: Skp2
C:Keywords: cell cycle control

Query Match 10.8%; Score 75; DB 2; Length 435;
Best Local Similarity 23.8%; Pred. No. 5.7;

Matches	40; Conservative	22; Mismatches	58; Indels	48; Gaps	8;
QY	4	PDNHRNKKVINDDDGNELGSGIME-LTDTETLI-----	LYTRKRDSVKVHYLCRR	54	
Db	97	PKLNRENFPGVSNLSLPDELLLGIFSCLCIPPELLKYSGVCKEYRLASDESIAQTLDLT-	155		
QY	55	YGYDNLNLFBSGRRCCOTGGGIFAFKCARA-----	SELFNMQLQEI----	94	
Db	156	---GKNLHPDVTGR--LLSQGVIAFRPSRFMDQPLAEHSPFPRQDMDLNSVIEVSTL	210		
QY	95	-----MQNNQINLV-VEEPVVVE-----RNNHQTELEVERTRTPTTP	129		
Db	211	HGATISQCKIKNLSLELRSDPEIWTIAKNSLVRLNLPCGCFPKFP	258		

RESULT 10
A47550
bride of sevenless precursor - fruit fly (*Drosophila virilis*)
N/Alternate names: boss
C/Species: *Drosophila virilis*
C/Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 21-Jul-2000
C/Accession: A47550
R/Hart, A.C.; Harrison, S.D.; Van Vactor Jr., D.L.; Rubin, G.M.; Zipursky, S.L.
Proc. Natl. Acad. Sci. U.S.A. 90, 5047-5051, 1993
A/Title: The interaction of bride of sevenless with sevenless is conserved between *Drosophila* and *Arabidopsis thaliana*
A/Reference number: A47550; MUID:93281693; PMID:8506350
A/Accession: A47550
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-893 <HAR>
A/Cross-references: CB:L08132; NID:G290215; PIDN:AAA72332.1; PID:G290216
C/Genetics:
A/Gene: FlyBase: Dvir/boss
A/Cross-references: FlyBase: FBgn0013105
A/Intons: 14/3; 175/3; 324/3; 809/3
A/Keywords: transmembrane protein

```

Query Match      10.5%; Score 73; DB 2; Length 893;
Best Local Similarity 24.1%; Pred.No. 23;
Matches 27; Conservative 19; Mismatches 50; Indels 16; Gaps 5;

Qy 20 GNELGSGIMELPDTLLIYTRKRSVKHWYLCRRGYDSN---LPSFSGRRCTGGGI 76
      G K I A S : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 253 GKIAS--VEIYETLTDIYNEFDAAQGEHMLFKFSDNVYILFGKQANHFKENGTV 310
      : : : : : : : : : : : : : : : : : : : : : : : : : :

Qy 77 FAFKCARBELFMNLQEIOMNNSINVVPEEPVFNRNHHQTELEVPTPTPT 128
      : : : : : : : : : : : : : : : : : : : : : : : : : :

Db 311 FAVPTERTPHIE--LEE-LPNKSTILMENG-----ELRVADLDPMPTT 351
      : : : : : : : : : : : : : : : : : : : : : : : : : :

```

RESULT 11

T37308

Affase homolog - Cryptosporidium parvum

C/Species: Cryptosporidium parvum

C/Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 18-Aug-2000

C/Accession: T37308

R/Zhu, G.; Keithly, J.S.

Mol. Biochem. Parasitol. 90, 307-316, 1997

A/Title: Molecular analysis of a P-type ATPase from *Cryptosporidium parvum*.

A/Reference number: Z21674; MUID:98156766; PMID:19497052

A/Accession: T37308

A/Status: preliminary; translated from GB/EMBL/DDBJ

A/Molecule type: DNA

A/Residues: 1-1528 <ZHU>

A/Cross-references: EMBL:U65981; NID:g1513209; PID:g1513210; PID:AAC47833.1

C/Genetics:

A/Note: CpA-E1

C/Superfamily: ATPase nucleotide-binding domain homology

P:986-1183/Domat: ATPase nucleotide-binding domain homology <ATN>

Query Match	10.4%	Score 72.5;	DB 2;	Length 1528;
Best Local Similarity	22.1%;	Pred. No. 50;		
Matches	27;	Conservative	16;	Mismatches
			46;	Indels
				33;
				Gaps
				4

Qy	8	RNKRVINVDGNGELSGIMELTDLEILYTRKRDVSKWVKHCLRYGVGDNGLNSFESG	67
Dd	567	RNTKTLNISFTGNQSFSFVEQS-----NSCVSFYPT	602
Qy	68	RCRCGGGIFAFKCAR---ABEFLMLOEQMNNSINVPEPVBERNNHQTELEVPRTER	124
Dd	603	RGSDPFGGIFLSSSLTONIKBEINNLI-HFNKNKTIENTENHELESN----FEMPDPDX	656
Qy	125	TP 126	
Dd	657	IP 658	

```

RESULT 12
T04499
auxin-induced protein homolog F8F16.140 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change
C:Accession: T04499
R:Sevan, M.; Brandt, P.; Dose, S.; Jarke, D.; Scharfe, M.; Schö
submitted to the Protein Sequence Database, April 1998
A:Reference number: Z15375
A:Accession: T04499
A:Molecule type: DNA
A:Residues: 1-189 <REV>
A:Cross-references: EMBL:AL021633
A:Experimental source: cultivar Columbia; BAC clone F8F16
C:Genetics:
A:Map position: 4
A:Note: F8F16.140

```

```

Query Match      10.3%; Score 72; DB 2; Length 189;
Best Local Similarity 24.2%; Prid. No. 4.2;
Matches          40; Conservative 21; Mismatches 52; Indels 52; Gaps 10;

Qy    1 DTVPDNHNKFKVINDDGNELGSG-----IMELTDTLEIL-----YTKR 42
      ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
Db    38 DNVDSDNN--ISININNNGSGSGSKIKFLKRTLSTDTATAPGYLAHSVCKEKR 95
      ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||

Qy    43 DSVKWHYL-----CLRRYGDNLFSFGRRCTQGCFIAFKCARAEELFNMLQEIMQ 96
      ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||

Db    96 YKIPTYELSHQA FVHLRAEE--CGFQA-----CILRIPEVA--VFESILKIME 144
      ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||

Qy    97 NNS-----INVVEPFVVERNNHQTELEVPTP-RTPTTP 129
      ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||

Db   145 DWKSDAVITOCPCENATSEVM---SYRHPSCPRTSPHOHSP 186
      ||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||

```

RESULT 13
D70313
riboflavin kinase - Aquifex aeolicus
C:Species: Aquifex aeolicus
C:Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 12-Jun-2003
C:Accession: D70313
R:Deckert, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; Ove
V.
Nature 392, 353-358, 1998
A:Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.
A:Reference number: A70300; MUID:95196666; PMID:9537320
A:Accession: D70313
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-314 <AQF>
A:Cross-references: GB:AE000675; NID:G2982863; PIDN:AAC06488.1; PID:G2982868; GB:AE000665;
A:Experimental source: strain VF5
C:Genetics:

Query Match	10.3%;	Score 72;	DB 2;	Length 314;
Best Local Similarity	25.5%;	Pred. No. 7.9;		
Matches 28;	Conservative 22;	Mismatches 34;	Indels 26;	Gaps 5

Search completed: August 2, 2004, 09:31:49
Job time : 20.6711 secs

Db 1658 ---VDIYKCKRLEHVFTSSMVGSLSQLQELHISNCSEMEEVIVKDADDSVEEDKEKESD 1713

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 2, 2004, 09:23:08 ; Search time 13.6842 Seconds
(without alignments)
494.666 Million cell updates/sec

Title: US-09-757-415a-1_COPY_11_140
Perfect score: 696
Sequence: 1 DTPVDNRHNRKFKVINDDGG.....NNHQTELEVPTPTPTTPG 130

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	145	20.8	482	1 DOK1_MOUSE	P97465 mus musculus
2	143.5	20.6	481	1 DOK1_HUMAN	O99704 homo sapien
3	135	19.4	412	1 DOK2_HUMAN	C60496 homo sapien
4	128	18.4	412	1 DOK2_MOUSE	O70469 mus musculus
5	124.5	17.9	306	1 DOK5_HUMAN	O99104 homo sapien
6	121.5	17.5	306	1 DOK5_MOUSE	O912m9 mus musculus
7	82	11.8	1235	1 IRS1_RAT	P35570 rattus norv
8	80.5	11.6	427	1 DPF3_CHICK	P58270 gallus gall
9	79.5	11.4	394	1 Z200_HUMAN	P98182 homo sapien
10	78	11.2	1233	1 IRS1_MOUSE	P35569 mus musculus
11	78	11.2	1242	1 IRS1_HUMAN	P35568 homo sapien
12	73.5	10.6	322	1 DUSC_VIBPA	O87m01 vibrio para
13	73	10.5	893	1 BOSS_DROVI	O24738 drosophila
14	72	10.3	535	1 Z257_HUMAN	O9y2q1 homo sapien
15	71.5	10.3	1309	1 RAD9_YEAST	P14737 saccharomyc
16	71	10.2	493	1 MUE1_OCEIH	O8c2c6 oceanobacill
17	70	10.1	275	1 TRUP_SVNY3	F73295 synechocyst
18	69.5	10.0	612	1 FEL_HPV1A	F03111 human papil
19	69	9.9	732	1 POK_DROME	O01842 drosophila
20	69	9.9	742	1 SYG_CAEEL	O10039 caenorhabdi
21	69	9.9	787	1 AGL2_BACTO	O9f234 bacillus th
22	68.5	9.8	655	1 LY9_HUMAN	O9hbq7 homo sapien
23	68.5	9.8	795	1 SYFE_SHEON	O8ef99 shewanella
24	68.5	9.8	1013	1 HGBE_HAEIN	O9kxv2 haemophilus
25	68.5	9.8	1565	1 PAC_STRVU	P11657 streptococc
26	68	9.8	608	1 LEPA_AGRTS	O9uig2 agrobacteri
27	68	9.8	1324	1 IRS2_HUMAN	O9y4h2 homo sapien
28	67	9.6	675	1 ZG20_XENLA	P18714 xenopus lae
29	67	9.6	852	1 RA50_THEMA	O9xi11 thermotoga
30	66.5	9.6	237	1 PLSC_HELPJ	O9znj8 helicobacte
31	66.5	9.6	240	1 PLSC_HELPJ	O25903 helicobacte
32	66.5	9.6	605	1 TC17_HUMAN	O60765 homo sapien
33	66.5	9.6	2555	1 PPS3_BACSU	P39847 bacillus su

ALIGNMENTS

RESULT 1
DOK1_MOUSE
ID DOK1_MOUSE STANDARD; PRT: 482 AA.
AC P97465: Q9R213;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Docking protein 1 (p62(dok)) (Downstream of tyrosine kinase 1).
GN DOK1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RC SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX STRAIN=C57BL/6;
RX MEDLINE=97160841; PubMed=9008161;
RA Yamanashi Y., Baltimore D.;
RT "Identification of the Abl- and rasGAP-associated 62 kDa protein as a docking protein, Dok.";
RL Cell 88:205-211(1997).
RN [2]
RP REVISIONS TO 381 AND 384.
RA Yamanashi Y., Baltimore D.;
RL Submitted (MAY-2002) to the EMBL/GenBank/DDBJ databases.
RN [3]
RC SEQUENCE FROM N.A.
RX STRAIN=129/SVJ;
RX MEDLINE=99126643; PubMed=9927484;
RA Jang W., Hua A., Spilson S.V., Miller W., Roe B.A., Meisler M.H.;
RT "Comparative sequence of the Abl- and mouse BAC clones from the mnd2 region of chromosome 2p13.";
RL Genome Res. 9:53-61(1999).
RN [4]
RC SEQUENCE FROM N.A.
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M.J., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Hellon E., Kettner M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grumwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

Q9aby6 caulobacter
Q92su3 rhizobium m
P40562 saccharomyc
Q03392 schizosacch
P97360 mus musculu
Q9byg5 homo sapien
O8ydb8 brucella me
Q8fv17 brucella su
P06829 sendai viru
Q58226 methanococc
Q9kt00 vibrio chol
P34316 caenorhabdi

34 66 9.5 552 1 SYK_CAUCR
35 66 9.5 608 1 LEPA_RHIME
36 66 9.5 993 1 YIS2_YEAST
37 65.5 9.4 260 1 PCNA_SCHPO
38 65.5 9.4 485 1 ETV6_MOUSE
39 65 9.3 372 1 PAGB_HUMAN
40 65 9.3 602 1 LEPA_BRUME
41 65 9.3 602 1 LEPA_BRUSU
42 64.5 9.3 2048 1 RRPL_SENDE
43 64 9.2 297 1 Y816_METJA
44 64 9.2 323 1 DUSC_VIBCH
45 64 9.2 358 1 YKT5_CAEEL


```
KW Phosphorylation; Alternative splicing.
FT DOMAIN 4 119 PH.
FT DOMAIN 152 254
FT MOD_RES 362 362 PHOSPHORYLATION (BY INSR).
FT MOD_RES 398 398 PHOSPHORYLATION (BY INSR).
FT VARSPLIC 153 177 SQWVTVQTEAAERCLGHSYVLR -> HVLFRGPPPLPL
RPNHLPLDGTGK (in isoform 2).
FT VARSPLIC 178 481 Missing (in isoform 2).
FT VARSPLIC 178 481 /FTID=VSP 003852.
FT MUTAGEN 362 362 Y-3F: NO ASSOCIATION WITH NCK.
FT MUTAGEN 398 398 Y-3F: NO ASSOCIATION WITH GAP; WHEN
FT MUTAGEN 398 398 ASSOCIATED WITH F-362.
FT CONFLICT 1 20 MDGAVMEGPIFLQSGRFGTK -> RLPAQASATREPRWS
PFO (IN REF. 3).
SQ SEQUENCE 481 AA; 52391 MW; E9D947831244BA6C CRC64;

Query Match 20.6%; Score 143.5; DB 1; Length 481;
Best Local Similarity 29.5%; Pred. No. 2e-07;
Matches 33; Conservative 20; Mismatches 50; Indels 9; Gaps 2;

QY 24 GSGIMELTDLEILYTRKRDSD-----VKWHLCLRRYGYDSNLFSPESGRRCTQGGIF 77
D 28-FEB-2003 (Rel. 41, Created)
D 28-FEB-2003 (Rel. 41, Last sequence update)
D 10-OCT-2003 (Rel. 42, Last annotation update)
D Docking protein 2 (p56(dok-2)) (Downstream of tyrosine kinase 2).
D DOK2.
GN Homo sapiens (Human).
OS Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98148015; PubMed=9478921;
RA Di Cristofano A., Carpino N., Dunant N., Friedland G., Kobayashi R.,
RA Strife A., Wisniewski D., Clarkson B., Pandolfi P.P., Resh M.D.;
RT "Molecular cloning and characterization of p56dok-2 defines a new
RT family of RasGAP-binding proteins.";
RL J. Biol. Chem. 273:4827-4830(1998).
CC -1- TISSUE SPECIFICITY: Expressed in peripheral blood leukocytes,
CC and mediate particular biological responses.
CC -1- SUBUNIT: Interacts with RET (By similarity). Interacts with
CC RasGAP.
CC -1- TISSUE SPECIFICITY: Expressed in peripheral blood leukocytes,
CC lymph nodes and spleen.
CC -1- DOMAIN: PTB domain mediates receptor interaction.
CC -1- PTM: Tyrosine phosphorylated.
CC -1- SIMILARITY: BELONGS TO THE DOK FAMILY.
CC -1- SIMILARITY: Contains 1 PH domain.
CC -1- SIMILARITY: Contains 1 PH domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF034970; AAC13265.1; -
DR Genbank; HGNC:2991; DOK2.
```

```
DR MIM; 604997; -.
DR GO; GO:0005515; P:protein binding; TAS.
DR GO; GO:0007166; P:cell surface receptor linked signal transdu. .; TAS.
DR InterPro; IPR002404; Insn_receptorSI.
DR InterPro; IPR001849; PH.
DR Pfam; PF02174; IRS; 1.
DR Pfam; PF00169; PH; 1.
DR SMART; SM00233; PH; 1.
DR SMART; SM00310; PTBI; 1.
DR PROSITE; PS50003; PH_DOMAIN; 1.
KW Phosphorylation.
FT DOMAIN 4 114
FT DOMAIN 148 247
FT DOMAIN 247 247
SQ SEQUENCE 412 AA; 45548 MW; A8B08B07F0DA91EC CRC64;

Query Match 19.4%; Score 135; DB 1; Length 412;
Best Local Similarity 34.0%; Pred. No. 1.3e-06;
Matches 33; Conservative 13; Mismatches 33; Indels 18; Gaps 3;

QY 47 WHYLCRLRYGYDSNLFSPESGRRCTQGGIFAFKCARAEELFNLMQEIIM--QNNSINVVE 104
D 28-FEB-2003 (Rel. 41, Last sequence update)
D 10-OCT-2003 (Rel. 42, Last annotation update)
D Docking protein 2 (p56(dok-2)) (Downstream of tyrosine kinase 2) (Dok-
D related protein) (Dok-R) (IL-four receptor interacting protein)
D (FRIP).
GN DOK2 OR FRIP.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98361257; PubMed=9697832;
RA Nelms K., Snow A.L., Hu-Li J., Paul W.E.;
RT "FRIP, a hematopoietic cell-specific rasGAP-interacting protein
RT phosphorylated in response to cytokine stimulation.";
RL Immunity 9:13-24(1998).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=98148015; PubMed=9478921;
RA Di Cristofano A., Carpino N., Dunant N., Friedland G., Kobayashi R.,
RA Strife A., Wisniewski D., Clarkson B., Pandolfi P.P., Resh M.D.;
RT "Molecular cloning and characterization of p56dok-2 defines a new
RT family of RasGAP-binding proteins.";
RL J. Biol. Chem. 273:4827-4830(1998).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=98435849; PubMed=9764820;
RA Jones N., Dumont D.J.;
RT "The Tek/Tie2 receptor signals through a novel Dok-related docking
RT protein, Dok-R.";
RL Oncogene 17:1097-1108(1998).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=23388257; PubMed=12477932;
RA Strausberg R.L., Feingold B.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
```


RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren B.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettaman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalilus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [5].
RP SEQUENCE OF 42-306 FROM N.A. (ISOFORM 1).
RC TISSUE=Kidney;
RA Ottenwaelder B., Obermaier B., Mewes H.-W., Gassenhuber J.,
RA Wiemann S.;
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Docking proteins interact with receptor tyrosine kinases
CC and mediate particular biological responses. Functions in RET-
CC mediated neurite outgrowth and plays a positive role in activation
CC of the MAP kinase pathway. Putative link with downstream effectors
CC of RET in neuronal differentiation (By similarity).
CC -!- SUBUNIT: Interacts with phosphorylated RET. In contrast to other
CC DOK proteins, it does not interact with RASGAP (By similarity).
CC -!- ALTERNATIVE PRODUCTS:
CC Event-Alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoId=Q9P104-1; Sequence=Displayed;
CC IsoId=Q9P104-2; Sequence=VSP_003854;
CC Note=No experimental confirmation available;
CC -!- DOMAIN: PTB domain mediates receptor interaction (By similarity).
CC -!- SIMILARITY: BELONGS TO THE DOK FAMILY.
CC -!- SIMILARITY: Contains 1 PTB domain.
CC -!- SIMILARITY: Contains 1 PTB domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF132732; AAF66443.1; -;
DR EMBL; AF463268; AAL74194.1; -;
DR EMBL; AL118501; CAC00469.1; -;
DR EMBL; AL162292; CAC32860.1; -;
DR EMBL; BC008992; AAH08992.1; -;
DR EMBL; AL050069; CAB43255.1; -;
DR PIR; T08731; T08731.
DR Genew; HGNC:16173; DOK5.
DR InterPro; IPR02404; Insln_receptorS1.
DR InterPro; IPR001849; PH.
DR Pfam; PF02174; IRS; 1.
DR Pfam; PF00169; PH; 1.
DR SMART; SM00233; PH; 1.
DR SMART; SM00310; PTB; 1.
DR PROSITE; PDS0003; PH_DOMAIN; FALSE_NEG.
KW Alternative splicing.
FT DOMAIN 8 112 PH.
FT DOMAIN 130 232 PTB.
FT VARSPLOC 1 108 Missing (in isoform 2).
FT FTID=VSP_003854.
FT SAAIATAE0H -> LLQMKMSRA (IN REF. 1).
FT CONFLICT 225 234
SEQUENCE 306 AA; 35463 MW; 2F259529EB068DBE CRG64;
Query Match 17.9%; Score 124.5; DB 1; Length 306;
Best Local Similarity 26.3%; Pred. No. 1.1e-05;
Matches 30; Conservative 29; Mismatches 40; Indels 15; Gaps 3;
QY 24 GSGIMELTDTELLYTRKDSVK--WHYLCLRRYGYDSNLFSGRRCQTGGIFAPK 80

Db 153 GECALQITVEYICLDWQNVKPLSLGALRYGRDITTFPEAGMCTGSLFIQ 212
QY 81 CARAEELFNML-----QEIMONNSINVVEEPVWER-NNHQTELEVPRT 122
Db 213 TRDEAIVQKVHSAALAAIEAHERLLQSVKMLQMKNSRAASLSTWVPLRS 266
RESULT 6
DOK5 MOUSE
ID DOK5 MOUSE STANDARD; PRT; 306 AA.
AC Q91ZM9; O3CSM5;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Docking protein 5 (Downstream of tyrosine kinase 5) (Protein dok-5).
GN DOK5.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A., TISSUE SPECIFICITY, FUNCTION, AND INTERACTION WITH
RP RET.
RC TISSUE=Embryo;
EX MEDLINE=21363571; PubMed=11470823;
RA Grimm J., Sachs M., Britsch S., Di Cesare S., Schwarz-Romond T.,
RA Alltalo K., Birchmeier W.;
RT "Novel p62dok family members, dok-4 and dok-5, are substrates of the
RT c-ret receptor tyrosine kinase and mediate neuronal
RT differentiation";
RL J. Cell Biol. 154:345-354(2001).
RN [2]
RP SEQUENCE OF 4-306 FROM N.A.
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Aikawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojohori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Stauber F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection";
RL Nature 409:685-690(2001).
CC -!- FUNCTION: Docking proteins interact with receptor tyrosine kinases
CC and mediate particular biological responses. Functions in RET-
CC mediated neurite outgrowth and plays a positive role in activation
CC of the MAP kinase pathway. Putative link with downstream effectors
CC of RET in neuronal differentiation.
CC -!- SUBUNIT: Interacts with phosphorylated RET. In contrast to other
CC DOK proteins, it does not interact with RASGAP.
CC -!- TISSUE SPECIFICITY: Specifically expressed in the brain, with a
CC high specificity for neurons.
CC -!- DEVELOPMENTAL STAGE: In E12.5 and E13 embryos, it is expressed in
CC the central nervous system, e.g., in the neural tube, the dorsal
CC root and the cranial ganglion.
CC -!- DOMAIN: PTB domain mediates receptor interaction.
CC -!- SIMILARITY: BELONGS TO THE DOK FAMILY.
CC -!- SIMILARITY: Contains 1 PH domain.
CC -!- SIMILARITY: Contains 1 PTB domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL; AF418208; AAL14627.1; -
 CC EMBL; AK012430; BAE28233.1; -
 CC MGD; MGI:1924079; Dok5.
 CC GO; GO:0005066; F:transmembrane receptor protein tyrosine kin. . .; IPI.
 CC GO; GO:0001655; P:MAPKK cascade; IDA.
 CC GO; GO:0007399; P:neurogenesis; IDA.
 CC GO; GO:0007169; P:transmembrane receptor protein tyrosine kin. . .; IPI.
 CC InterPro; IPR002404; Insln_receptorSl.
 CC InterPro; IPR001849; PH.
 CC Pfam; PF02174; IRS; 1.
 CC SMART; SM00233; PH; 1.
 CC SMART; SM00310; PTBI; 1.
 CC PROSITE; PS50003; PH_DOMAIN; FALSE_NEG.
 CC DOMAIN 8 112 PH.
 CC DOMAIN 130 232 PTB.
 CC SEQUENCE 306 AA; 35453 MW; 7889802PBEAC04A6 CRC64;

Query Match 17.58; Score 121.5; DB 1; Length 306;
 Best Local Similarity 26.38; Pred. No. 2.3e-05;
 Matches 30; Conservative 28; Mismatches 41; Indels 15; Gaps 3;

QY 24 GSGIMEYTDLELYTRKRDSVK--WHYLCLRRYGYDNLFFESGRRCQTGGQIFAFK 80
 DB 153 GECALQITYEICLDVQNRVLKSLWPLSALRYGRDITWTFEAGRVCETGEGFIQ 212

QY 81 CARAEELFNML-----GEIQNNNSINVEEPVVER-NHOTELEVPRT 122
 DB 213 TRDGEAIYQVHSAALAEQHERLLQSVKNSMLQMKKGERAASLTSTVYPLPRS 266

RESULT 7
 ID_IRS1_RAT STANDARD; PRT; 1235 AA.
 AC P35570;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Insulin receptor substrate-1.
 DE IRS1 OR IRS-1.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RC STRAIN-Sprague-Dawley, TISSUE=Liver;
 RX MEDLINE=91287824; PubMed=1648180;
 RA Sun X.J., Rothenberg P., Kahn C.R., Backer J.M., Araki E.,
 RA Wilden P.A., Cahill D.A., Goldstein B.J., White M.F.;
 RT "Structure of the insulin receptor substrate IRS-1 defines a unique
 RT signal transduction protein.";
 RL Nature 352:73-77(1991).
 RN [2]
 RP PHOSPHORYLATION SITES.
 RX MEDLINE=94067102; PubMed=7504175;
 RA Sun X.J., Crimmins D.L., Myers M.G., Miralpeix M., White M.F.;
 RT "Pleiotropic insulin signals are engaged by multisite phosphorylation
 RT of IRS-1.";
 RL Mol. Cell. Biol. 13:7418-7428(1993).
 RN [3]
 RP PHOSPHORYLATION SITES.
 RX MEDLINE=93352637; PubMed=8349691;
 RA Tanasitjevic M.J., Myers M.G., Thoma R.S., Crimmins D.L., White M.F.,
 RA Sacks D.B.;
 RT "Phosphorylation of the insulin receptor substrate IRS-1 by casein

RT kinase II.";
 RL J. Biol. Chem. 268:18157-18166(1993).
 CC FUNCTION: May mediate the control of various cellular processes by
 CC insulin. When phosphorylated by the insulin receptor binds
 CC specifically to various cellular proteins containing SH2 domains
 CC such as phosphatidylinositol 3-kinase p85 subunit or GRB2.
 CC SIMILARITY: Contains 1 PH domain.
 CC SIMILARITY: Contains 1 PTB domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL; X58375; CAA41264.1; -
 CC PIR; S16948; S16948.
 CC HSP; P35568; IIRS.
 CC InterPro; IPR002404; Insln_receptorSl.
 CC InterPro; IPR001849; PH.
 CC Pfam; PF02174; IRS; 1.
 CC Pfam; PF00169; PH; 1.
 CC PRINTS; PRO0628; INSULINRSI.
 CC SMART; SM00233; PH; 1.
 CC SMART; SM00310; PTBI; 1.
 CC PROSITE; PS50003; PH_DOMAIN; 1.
 KW Phosphorylation.

FT DOMAIN 12 115 PH.
 FT DOMAIN 152 262 PTB.
 FT DOMAIN 872 881 POLY-GLN.
 FT DOMAIN 1196 1200 POLY-PRO.
 FT MOD_RES 99 99 PHOSPHORYLATION (BY CK2).
 FT MOD_RES 460 460 PHOSPHORYLATION (BY INSR).
 FT MOD_RES 502 502 PHOSPHORYLATION (BY CK2).
 FT MOD_RES 608 608 PHOSPHORYLATION (BY INSR).
 FT MOD_RES 628 628 PHOSPHORYLATION (BY INSR).
 FT MOD_RES 895 895 PHOSPHORYLATION (BY INSR).
 FT MOD_RES 939 939 PHOSPHORYLATION (BY INSR).
 FT MOD_RES 987 987 PHOSPHORYLATION (BY INSR).
 FT MOD_RES 1172 1172 PHOSPHORYLATION (BY INSR).
 FT MOD_RES 1222 1222 PHOSPHORYLATION (BY INSR).
 SQ SEQUENCE 1235 AA; 131178 MW; A274BC7540CA85C5 CRC64;

Query Match 11.8%; Score 82; DB 1; Length 1235;
 Best Local Similarity 30.0%; Pred. No. 1.8;
 Matches 24; Conservative 7; Mismatches 21; Indels 28; Gaps 2;

QY 52 LRRGYDNLFFESGRRCQTGGQIFAFKARAEELFNMLQEIQNNNSINVEEPVVER 111
 DB 206 IRRGSHSENFIEVGRSAVTGGEFWMQ-----VDSVVAQN 243

QY 112 NHOTELEVPRT-----PRT 125
 DB 244 MHEILAEAMRAMSDEFPRPT 263

RESULT 8
 ID_DP3_CHICK STANDARD; PRT; 427 AA.
 AC P58270;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Zinc-finger protein DPf3 (cer-d4).
 GN DPf3 OR CERD4.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
 OC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]

SEQUENCE FROM N.A. (ISOFORMS 1; 2; 3 AND 4).
RX MEDLINE=2183813; PubMed=11845289;
RA Ninkina N.N., Mertsalov I.B., Kulikova D.A., Alimova-Kost M.V.,
RA Simonova O.B., Korochkin L.I., Kiselev S.L., Buchman V.L.;
RT "Cerc4, third member of the d4 gene family: expression and
RT organization of genomic locus";
RL Mamm. Genome 12:862-866(2001).
CC -!- FUNCTION: Probably involved in RNA transcription.
CC -!- SUBCELLULAR LOCATION: Nuclear and cytoplasmic (Potential).
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=4;
CC Name=1;
CC IsoId=P58270-1; Sequence=Displayed;
CC Name=2;
CC IsoId=P58270-2; Sequence=VSP_005614, VSP_005615;
CC Name=3;
CC IsoId=P58270-3; Sequence=VSP_005614, VSP_005616;
CC Name=4;
CC IsoId=P58270-4; Sequence=VSP_005614, VSP_005615, VSP_005616;
CC -!- SIMILARITY: Belongs to the requiem/DPF family.
CC -!- SIMILARITY: Contains 2 PHD-type zinc fingers.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF362754; AAK51968.1; -;
CC EMBL; AF362753; AAK51967.1; -;
CC EMBL; AF362756; AAK51970.1; -;
CC EMBL; AF362755; AAK51969.1; -;
CC InterPro; IPR007087; Znf_C2H2.
CC Pfam; PF00628; Znf_C2H2; 2.
CC Pfam; PF00096; Znf_C2H2; 1.
CC SMART; SM00249; PHD; 2.
CC SMART; SM00355; Znf_C2H2; 1.
CC PROSITE; PS01359; ZF_PHD_1; 1.
CC PROSITE; PS00016; ZF_PHD_2; 2.
CC PROSITE; PS00028; ZINC_FINGER_C2H2_1; 1.
CC PROSITE; PS0157; ZINC_FINGER_C2H2_2; 1.
CC Transcription regulation; Zinc-finger; Nuclear protein; Repeat;
KW Alternative splicing
FT ZN_FING 247 270 C2H2-TYPE.
FT ZN_FING 308 368 PHD-TYPE 1.
FT ZN_FING 365 415 PHD-TYPE 2.
FT VARSP_LIC 102 137 Missing (in isoform 2, isoform 3 and
FT isoform 4).
FT
FT VARSP_LIC 212 224 /FTId=VSP_005614.
FT Missing (in isoform 2 and isoform 4).
FT
FT VARSP_LIC 340 427 /FTId=VSP_005615.
FT GPTCQFTNTTEAVKTVQMCIECKSCSLGCTGENDDLQ
FT LFCDDCDRGHYHCLNPPVPEPPEGSWSCHLCRELLERAS
FT AFGQQA -> AHLGREGREDEAAPTETEDLEFGTSSESDTS
FT TFGFEDDAERPLSRGGCGGSPSADKGGC (in
FT isoform 3 and isoform 4).
FT /FTId=VSP_005616.
FT
SQ SEQUENCE 427 AA; 48930 MW; 522DB95D1E714152 CRC64;

Query Match 11.6%; Score 80.5; DB 1; Length 427;
Best Local Similarity 23.4%; Pred.No.0.72;
Matches 33; Conservative 26; Mismatches 45; Indels 37; Gaps 6;

QY 9 NKFQVNVDDGNGELSGTMELTDTLLI-----LYTKRDSVKNWHYLCLE 54
DB 91 SRLKLEIKFETSHLPGKTELITETETFTTMSVDLRFSLCKLYTSEVD-----LPLKK 144
QY 55 YGVDSNLFSESGRRQCTGGGFAFRCARAEELFNMLQEIOMON-----NSINVVEEPV 107
DB 145 DGFSTSTLEALLR---GEGIEKKMDTKEDPIQIQRVLENDENADVNDEEDLEEDI 201

QY 108 VERNNHQTELEVPR-TPRTPT 127
DB 202 PKRKQR-----PRGRPKTPT 216

RESULT 9
Z200 HUMAN
ID Z200 HUMAN STANDARD; PRT; 394 AA.
AC P98182; O15361;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Zinc finger protein 200.
GN ZNF200 OR ZNFMP.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_taxID=9606;
RN [1]
SEQUENCE FROM N.A.
RX MEDLINE=99005537; PubMed=9787081;
RA Deng Z., Centola M., Chen X., Sood R., Vedula A., Fischel-Ghodsian N.,
RA Kastner D.L.;
RT "Identification of two Kruppel-related zinc finger genes (ZNF200 and
RL ZNF210) from human chromosome 16p13.3";
RN Genomics 53:97-103(1998).
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=97434208; PubMed=9288094;
RA Bernot A., Clepet C., Dasilva C., Devaud C., Petit J.-L.,
RA Caloustian C., Cruaud C., Samson D., Pulcini F., Weissenbach J.,
RA Heilig R., Notaricola C., Domingo C., Rozenbaum M., Bencherit E.,
RA Topaloglu R., Dewalle M., Dross C., Hadjari P., Dupont M.,
RA Demaille J., Touitou I., Smacou N., Nedelec B., Mery J.-P.,
RA Chaabouni H., Delpech M., Grateau G.;
RT "A candidate gene for familial Mediterranean fever";
RL Nat. Genet. 17:25-31(1997).
CC -!- FUNCTION: COULD HAVE A ROLE IN SPERMATOGENESIS.
CC -!- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -!- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN TESTIS, WEAKLY EXPRESSED
CC IN SPLEEN, THYMUS, PROSTATE, OVARY, SMALL INTESTINE COLON AND
CC PERIPHERAL BLOOD LEUKOCYTES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF060866; AAC70008.1; -;
CC EMBL; Y14443; CAA74795.1; ALT_INIT.
CC EMBL; AJ003147; CAA05905.1; ALT_INIT.
CC HSSP; P08046; 1A1I.
CC Genew; HGNC:12993; ZNF200.
CC MIM; 603231; -;
CC GO; GO:0005634; C:nucleus; NAS.
CC InterPro; IPR007087; Znf_C2H2.
CC Pfam; PF00096; Zf_C2H2; 5.
CC ProDom; PD000003; Znf_C2H2; 2.
CC SMART; SM00355; Znf_C2H2; 5.
CC PROSITE; PS00028; ZINC_FINGER_C2H2_1; 5.
CC PROSITE; PS0157; ZINC_FINGER_C2H2_2; 5.
CC Transcription regulation; Zinc-finger; Metal-binding; Nuclear protein.
FT ZN_FING 251 273 C2H2-TYPE.
FT ZN_FING 279 301 C2H2-TYPE.
FT ZN_FING 307 329 C2H2-TYPE.
FT ZN_FING 335 357 C2H2-TYPE.
FT ZN_FING 363 385 C2H2-TYPE.
FT CONFLICT 113 113 MISSING (IN REF. 2).

```
SO SEQUENCE 394 AA; 45402 MW; 8EB613P0F5E23377 CRC64;
Query Match 11.4%; Score 79.5; DB 1; Length 394;
Best Local Similarity 20.8%; Pred. No. 0.83;
Matches 27; Conservative 32; Mismatches 48; Indels 23; Gaps 6;
Qy 2 TVPDNHRNKEVNVDDGNGELSGMELT---DTELLYTRKRDVKVHY--LCRRYVG 56
Db 207 SIFQKRQRLVLTIND-----TPLELSKYVDISIIALTNRNRRRYTCLQCKQFN 261
Qy 57 YDNLRFESGRCQTGGQIFAFKCARAELEFNMQLQINQNSINNVVEPV-----V 108
Db 262 ESSVLISHQ---RTHIGEX--PYDCHHGKGFNFHKTNLNKHRIHTGERPYSCQCGKNF 316
Qy 109 ERNNHQTELE 118
Db 317 QNSHRSRHE 326
RESULT 10
IRS1_MOUSE STANDARD; PRT; 1233 AA.
AC P35569;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Insulin receptor substrate-1.
GN IRS1 OR IRS-1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
EN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94220494; PubMed=8167159;
RA Araki E., Haag B.L. III, Kahn C.R.;
RT "Cloning of the mouse insulin receptor substrate-1 (IRS-1) gene and
RT complete sequence of mouse IRS-1."
RL Biochim. Biophys. Acta 1221:353-356 (1994).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=94220494; PubMed=8448209;
RA Keller S.R., Abersold R., Garner C.W., Lienhard G.B.;
RT "The insulin-elicited 160 kDa phosphotyrosine protein in mouse
RT adipocytes is an insulin receptor substrate 1: identification by
RT cloning."
RL Biochim. Biophys. Acta 1172:323-326 (1993).
CC -|- FUNCTION: May mediate the control of various cellular processes by
CC insulin. When phosphorylated by the insulin receptor binds
CC specifically to various cellular proteins containing SH2 domains
CC such as phosphatidylinositol 3-kinase p85 subunit or GRB2.
CC -|- SIMILARITY: Contains 1 PH domain.
CC -|- SIMILARITY: Contains 1 PPS domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L24563; AAA39335.1; -
DR EMBL; G69722; CAA49378.1; -
DR PIR; S30185; S30185.
DR PDB; 1AYB; 31-AUG-94.
DR MGD; MGI:99454; IRS1.
DR InterPro; IPR002404; Insln_receptorS1.
DR InterPro; IPR001849; PH.
DR Pfam; PF00174; IRS; 1.
DR Pfam; PF00169; PH; 1.
DR PRINTS; PRO0628; INSULINRS1.
DR SMART; SMO0233; PH; 1.
```

```
DR SMART; SMO0310; PTBI; 1.
DR PROSITE; PS00003; PH_DOMAIN; 1.
KW Phosphorylation; 3D-Structure.
FT DOMAIN 12 115
FT DOMAIN 152 262
FT DOMAIN 675 680
FT DOMAIN 872 877
FT DOMAIN 1119 1128
FT DOMAIN 1194 1198
FT MOD_RES 99 99
FT MOD_RES 460 460
FT MOD_RES 608 608
FT MOD_RES 628 628
FT MOD_RES 891 891
FT MOD_RES 935 935
FT MOD_RES 983 983
FT MOD_RES 1173 1173
FT MOD_RES 1220 1220
FT CONFLICT 1038 1039
FT CONFLICT 1182 1182
FT CONFLICT 1233 AA; 130723 MW; C059B2D890DADD87 CRC64;
SQ SEQUENCE 1233 AA; 11.2%; Score 78; DB 1; Length 1233;
Query Match 11.2%; Score 78; DB 1; Length 1233;
Best Local Similarity 30.0%; Pred. No. 4.8;
Matches 21; Conservative 7; Mismatches 20; Indels 22; Gaps 1;
Qy 52 LRRYGVSNLFPESGRCQTGGQIFAFKCARAELEFNMQLQINQNSINNVVEPVVERN 111
Db 206 IRRGCHSENFFIEVGRSAVTGPGFWMQ-----VDDSVVAQN 243
Qy 112 NHOETELEVPR 121
Db 244 MHEILEAMR 253
RESULT 11
IRS1_HUMAN STANDARD; PRT; 1242 AA.
AC P35568;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Insulin receptor substrate-1 (IRS-1).
GN IRS1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
EN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93292738; PubMed=8513971;
RA Araki E., Sun X.J., Haag B.L. III, Chuang L.M., Zhang Y.,
RA Yang-Feng T.L., White M.F., Kahn C.R.;
RT "Human skeletal muscle insulin receptor substrate-1. Characterization
RT of the cDNA, gene, and chromosomal localization."
RL Diabetes 42:1041-1054 (1993).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=92181456; PubMed=1311924;
RA Nishiyama M., Wards J.R.;
RT "Cloning and increased expression of an insulin receptor substrate-1-
RT like gene in human hepatocellular carcinoma."
RL Biochem. Biophys. Res. Commun. 183:280-285 (1992).
```



```
FT TRANSMEM 752 772 POTENTIAL.
FT DOMAIN 773 893 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 183 307 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 307 337 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 328 328 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 471 471 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 482 482 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 893 AA; 99939 MW; AB99D4C21DB680C3 CRC64;

Query Match 10.5%; Score 73; DB 1; Length 893;
Best Local Similarity 24.1%; Pred. No. 11;
Matches 27; Conservative 19; Mismatches 50; Indels 16; Gaps 5;

QY 20 GNELGSGIMELDTTEILLTRKDSVKVHYLCIRRYGYDSN---LFSFESGRRCOTGGGI 76
DB 253 GKIAS--VEIVTETDIYNEFDARQRMHFKSDDNVYILFGNNMANHFKENGTV 310
QY 77 FAFKCARAEELNMLQEIOMNSINVVPRVVERNNHQTELEVPRTPPTT 128
DB 311 FAVPTERTSHIP--LEE-LPNKSFILMENG-----ELRTADLDNPTT 351

RESULT 14
2257 HUMAN
ID 2257 HUMAN STANDARD; PRT; 535 AA.
AC 09Y201;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Zinc finger protein 257 (Bone marrow zinc finger 4) (BMZF-4).
GN ZNF257 OR BMZF4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Bone marrow;
RX MEDLINE=20054457; PubMed=10585455;
RA Han Z.-G., Zhang Q.-H., Ye M., Kan L.-X., Gu B.-W., He K.-L.,
RA Shi S.-L., Zhou J., Fu G., Mao M., Chen S.-J., Yu L., Chen Z.;
RT "Molecular cloning of six novel Kruppel-like zinc finger genes from
RT hematopoietic cells and identification of a novel transregulatory
RT domain KRNB."
RL J. Biol. Chem. 274:35741-35748(1999).
CC -!- FUNCTION: May function as a transcription factor.
CC -!- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -!- SIMILARITY: BELONGS TO THE KRUEPPEL FAMILY OF C2H2-TYPE ZINC-
CC FINGER PROTEINS.
CC -!- SIMILARITY: Contains 1 KRAB domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF070651; AAD20957.1; -
CC HSSP; P08047; ISP2.
CC Genew; HGNC:13498; ZNF257.
CC MIM; 606957; -
CC GO; GO:0005634; C:nucleus; NAS.
CC GO; GO:0003677; F:DNA binding; NAS.
CC GO; GO:0008270; F:zinc ion binding; NAS.
CC GO; GO:0006355; P:regulation of transcription, DNA-dependent; NAS.
CC InterPro; IPR001909; KRAB.
CC InterPro; IPR007087; Znf_C2H2.
CC InterPro; IPR007086; Znf_C2H2_sub.
CC Pfam; PF001352; KRAB; 1.
CC Pfam; PF00096; zf-C2H2; 10.
CC PRINTS; PR00048; ZINCFINGER.
```

```
DR PRODOM; PD000003; Znf_C2H2; 9.
DR SMART; SM00349; KRAB; 1.
DR SMART; SM00355; Znf_C2H2; 12.
DR PROSITE; PS0805; KRAB; 1.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 10.
DR PROSITE; PS0157; ZINC_FINGER_C2H2_2; 12.
KW Transcription regulation; DNA-binding; Zinc-finger; Metal-binding;
KW Nuclear protein; Repeat.
FT DOMAIN 4 75 KRAB.
FT ZN_FING 173 195 C2H2-TYPE.
FT ZN_FING 201 223 C2H2-TYPE.
FT ZN_FING 229 251 C2H2-TYPE.
FT ZN_FING 257 282 C2H2-TYPE (DEGENERATE).
FT ZN_FING 288 310 C2H2-TYPE.
FT ZN_FING 316 338 C2H2-TYPE.
FT ZN_FING 344 366 C2H2-TYPE.
FT ZN_FING 372 397 C2H2-TYPE.
FT ZN_FING 403 425 C2H2-TYPE.
FT ZN_FING 431 453 C2H2-TYPE.
FT ZN_FING 459 481 C2H2-TYPE.
FT ZN_FING 487 509 C2H2-TYPE.
SQ SEQUENCE 535 AA; 62348 MW; 22DCSB0C4613BC51 CRC64;

Query Match 10.3%; Score 72; DB 1; Length 535;
Best Local Similarity 21.4%; Pred. No. 7.4;
Matches 18; Conservative 18; Mismatches 38; Indels 10; Gaps 2;

QY 51 CLRRYGYDSNLFSESGRRRCOTGGIGIFAFKCARAEELNMLQEIOMNSINVVPR--- 106
DB 262 CCKAFKWSALLTITQHKRIHTGK--PYKCECGKAFNOSSALTRHKMHTGKPFQCE 319
QY 107 ---VVERNNHQTELEVPRTPPT 126
DB 320 ECGKAFNSSLTQHKIHTKEP 343

RESULT 15
RAD9 YEAST
ID RAD9 YEAST STANDARD; PRT; 1309 AA.
AC P14737; Q04920;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE DNA repair protein RAD9.
GN RAD9 OR YDR217C OR YD9934.02C.
OS Saccharomyces cerevisiae (Baker's Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89313732; PubMed=2664461;
RA Schiestl R.H., Reynolds P., Prakash S., Prakash L.;
RT "Cloning and sequence analysis of the Saccharomyces cerevisiae RAD9
RT gene and further evidence that its product is required for cell cycle
RT arrest induced by DNA damage."
RL Mol. Cell. Biol. 9:1882-1896(1989).
RN [2]
RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
RX MEDLINE=91061763; PubMed=2247073;
RA Weinert T.A., Hartwell L.H.;
RT "Characterization of RAD9 of Saccharomyces cerevisiae and evidence
RT that its function acts posttranslationally in cell cycle arrest after
RT DNA damage."
RL Mol. Cell. Biol. 10:6554-6564(1990).
RN [3]
RP SEQUENCE FROM N.A.
RA Murphy L., Harris D., Barrell B.G., Rajadream M.A.;
RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.
RN [4]
RP PHOSPHORYLATION, AND INTERACTION WITH RAD53.
RX MEDLINE=98429491; PubMed=9755168;
RA Vialard J.E., Gilbert C.S., Green C.M., Lowndes N.F.;
```

```

RT "The budding yeast Rad9 checkpoint protein is subjected to
RT Mec1/Tell-dependent hyperphosphorylation and interacts with Rad53
RT after DNA damage."
RL EMOB J. 17:5679-5688(1998).
CC -!- FUNCTION: Essential for cell cycle arrest at the G2 stage
CC following DNA damage by X-irradiation or inactivation of DNA
CC ligase.
CC
CC -!- SUBUNIT: Physically associates with RAD53.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- SIMILARITY: Contains 1 BRCT domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; M25049; AAA34954.1; --
CC EMBL; Z48612; CAA88497.1; --
CC PIR; S59424; BV3YD9.
CC PDB; 1FHR; 18-OCT-00.
CC PDB; 1J4K; 05-DEC-01.
CC PDB; 1J4L; 05-DEC-01.
CC PDB; 1J4P; 05-DEC-01.
CC PDB; 1J4Q; 05-DEC-01.
CC PDB; 1K2M; 05-DEC-01.
CC PDB; 1K2N; 05-DEC-01.
CC PDB; 1K3N; 05-DEC-01.
CC PDB; 1K3O; 05-DEC-01.
CC Germline; 140709; --
CC SGD; S0002625; RAD9.
CC GO; GO:0005634; C:nucleus, IC.
CC GO; GO:0005515; F:protein binding; IDA.
CC GO; GO:0000077; P:DNA damage response, signal transduction re...; IMP.
CC GO; GO:0006289; P:nucleotide-excision repair; IMP.
CC GO; GO:0045944; P:positive regulation of transcription from P...; IMP.
CC GO; GO:0000074; P:regulation of cell cycle; IGI.
CC InterPro; IPR001357; BRCT.
CC Pfam; PF00533; BRCT; 1.
CC SMART; SM00292; BRCT; 1.
CC PROSITE; PS0172; BRCT; 1.
CC Cell cycle; DNA damage; DNA replication inhibitor; Nuclear protein;
CC Phosphorylation; 3D-structure.
CC DOMAIN 994 1122 BRCT.
CC CONFLICT 433 433 C->S (IN REF. 3).
CC SEQUENCE 1309 AA; 148413 MW; 6B77D39A95021F84 CRC64;

Query Match 10.3%; Score 71.5; DB 1; Length 1309;
Best Local Similarity 23.3%; Pred. No. 25;
Matches 38; Conservative 17; Mismatches 51; Indels 57; Gaps 7;

QY 13 VINVDDGNGELGSGIMELDTTELTYT-----RKDSVKWHVLCLEPYGYD----- 58
Db 334 VNNLDIDINAVSGTPTSRNABEEMYSSESVNNRFPKXWIF----RYSKDTENNNSRS 389
QY 59 -----SNLFFSGRCRCQT-----CQGIFAFK-----C 81
Db 390 TQIVNNPRTQEMPLDSISIDTQLSKSFNTTNELETQIVSCLSQGISAQKGPVVFHST 449
QY 82 ARAEELFMLOEIMQNNINNV-EEPV-VERNNHQTELEVPT 122
Db 450 QTTEEINTQINSPEQNALNATFETVLSRINFEPILFVPT 492

```

Search completed: August 2, 2004, 09:29:56
Job time : 14.6842 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 2, 2004, 09:26:19 ; Search time 58.1579 Seconds
(without alignments)
705.276 Million cell updates/sec

Title: US-09-757-415a-1_copy_11_140

Perfect score: 696

Sequence: 1 DTVPDNRHKNFKVINDDG.....NNHQTELEVPRTPTPTTPG 130

Scoring table: BLOSUM62

Gapop 10.0 ; Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_25:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phase:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_rvirus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	696	100.0	512	4	Q8WU20 homo sapien
2	685	98.4	508	11	Q8C180 mus musculus
3	616.5	88.6	509	13	Q90ZF5 xenopus lae
4	616.5	88.6	509	13	Q7ZWM2 xenopus lae
5	613.5	88.1	509	13	Q8UVU3 xenopus lae
6	510.5	73.3	492	4	Q43559 homo sapien
7	484.5	69.6	492	11	Q91WJ0 mus musculus
8	178	25.6	195	5	Q44830 caenorhabdi
9	174	25.0	442	5	Q8VLL4 drosophila
10	174	25.0	442	5	Q8MT60 drosophila
11	141	20.3	444	11	Q9QZK7 mus musculus
12	137.5	19.8	223	4	Q8WU28 homo sapien
13	137.5	19.8	223	4	Q86LJ2 homo sapien
14	137.5	19.8	257	4	Q86H12 homo sapien
15	135	19.4	412	4	Q8N544 homo sapien
16	134.5	19.3	622	5	Q9W3R6 drosophila

17	124.5	17.9	311	4	Q75209	Q75209 homo sapien
18	124.5	17.9	326	4	Q8TEW6	Q8TEW6 homo sapien
19	124.5	17.9	326	4	Q8ETP2	Q8ETP2 homo sapien
20	124.5	17.9	326	4	Q8NVV3	Q8NVV3 homo sapien
21	121.5	17.5	306	11	Q8BR13	Q8BR13 mus musculus
22	120.5	17.3	325	11	Q99KE3	Q99KE3 mus musculus
23	120.5	17.3	496	4	Q9H666	Q9H666 homo sapien
24	114	16.4	320	13	Q7ZX79	Q7ZX79 xenopus lae
25	85	12.2	1807	5	Q81389	Q81389 plasmodium
26	84	12.1	598	10	Q8LU94	Q8LU94 arabidopsis
27	83	11.9	570	11	Q8C8U7	Q8C8U7 mus musculus
28	80.5	11.6	777	5	Q9VL91	Q9VL91 drosophila
29	80.5	11.6	777	5	Q44122	Q44122 drosophila
30	79.5	11.4	393	4	Q7ZSV1	Q7ZSV1 homo sapien
31	79	11.4	549	5	Q9U3B6	Q9U3B6 caenorhabdi
32	79	11.4	1837	5	Q86905	Q86905 dictyosteli
33	78.5	11.3	475	16	Q7VEX6	Q7VEX6 helicobacte
34	78	11.2	1251	6	Q82224	Q82224 cercopithe
35	77	11.1	4261	5	Q8IFP4	Q8IFP4 plasmodium
36	76.5	11.0	1913	5	Q9GR96	Q9GR96 leucophaea
37	76	10.9	431	10	Q9FIQ7	Q9FIQ7 arabidopsis
38	76	10.9	507	16	Q7VBX2	Q7VBX2 prochloroco
39	76	10.9	914	2	Q9AQF3	Q9AQF3 ciostroidium
40	75.5	10.8	320	10	Q9M006	Q9M006 arabidopsis
41	75.5	10.8	753	10	Q9C7W5	Q9C7W5 arabidopsis
42	75.5	10.8	1025	4	Q8NDL1	Q8NDL1 homo sapien
43	75.5	10.8	1114	4	Q9H0F1	Q9H0F1 homo sapien
44	75.5	10.8	1354	4	Q75183	Q75183 homo sapien
45	75	10.8	537	6	Q9WZU5	Q9WZU5 sus scrofa

ALIGNMENTS

RESULT 1

Q8WU20 PRELIMINARY; PRT; 512 AA.

AC Q8WU20; Q43558; (TREMBLrel. 20, Created)
 DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE Suct-associated neurotrophic factor target (FGFR signalling adaptor
 DE SNT-1).
 GN SNT-1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=uterus;
 RA Strausberg R.;
 RL Submitted (JAN-2002) to the EMBL/GenBank/DDBJ databases.
 RN [2]
 RP SEQUENCE OF 1-508 FROM N.A.
 RC MEDLINE=98324992; PubMed=9660748;
 RX Xu H., Lee K.W., Goldfarb M.;
 RA "Novel recognition motif on fibroblast growth factor receptor mediates
 RT direct association and activation of SNT adaptor proteins.";
 RL J. Biol. Chem. 273:17987-17990(1998).
 DR EMBL; BC021562; AAH21562.1;
 DR EMBL; AF036717; AAB92554.1;
 DR Genew; HGNC:16971; FR52.
 DR GO; GO:0005624; C:membrane fraction; TAS.
 DR GO; GO:0005158; F:insulin receptor binding; IEA.
 DR GO; GO:0008543; P:FGF receptor signaling pathway; TAS.
 DR GO; GO:0007165; P:signal transduction; TAS.
 DR InterPro; IPR002404; Insln_receptorS1.
 DR Pfam; PF02174; IRS; 1.
 DR SMART; SM00310; PTBI; 1.
 SQ SEQUENCE 512 AA; 57527 MW; 7B4F07D1B33714EE CRC64;

```

Query Match      100.0%; Score 696; DB 4; Length 512;
Best Local Similarity 100.0%; Pred. No. 4.5e-68;
Matches 130; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DTVPDNRHNRKFKVINVDGNGELSGIMELTDTELLYTRKRDSSVKWHYLCRLRYGYDSN 60
DB 11 DTVPDNRHNRKFKVINVDGNGELSGIMELTDTELLYTRKRDSSVKWHYLCRLRYGYDSN 70
QY 61 LFSFESGRCCTGGGIFAFKCARABELFNMLQEIQNNSINNVVEPVPVVRNNHOTELEVP 120
DB 71 LFSFESGRCCTGGGIFAFKCARABELFNMLQEIQNNSINNVVEPVPVVRNNHOTELEVP 130
QY 121 RTPRTPTTPG 130
DB 131 RTPRTPTTPG 140

RESULT 2
Q8C180 PRELIMINARY; PRT; 508 AA.
AC Q8C180;
DT 01-MAR-2003 (TREMELrel. 23, Created)
DT 01-MAR-2003 (TREMELrel. 23, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE SUC1-associated neurotrophic factor target (Hypothetical protein)
OS Mus musculus (Mouse)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Skin;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA "Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs."
RL Nature 420:563-573(2002).
[2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Brain;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Joquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smallus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RA "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences."
Proc. Natl. Acad. Sci. U.S.A. 99:16999-16903(2002).
[3]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Brain;
RA Strausberg R.;
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
[4]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Brain;
RA Strausberg R.;
RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.

```

```

DR EMBL; AK029813; BAC26132.1; -
DR EMBL; BC043109; AAH43109.1; -
DR EMBL; BC055334; AAH55334.1; -
DR GO; GO:0005158; F:Insulin receptor binding; IEA.
DR InterPro; IPR002404; Insln_receptors1.
DR Pfam; PF02174; IRS; 1.
DR SMART; SM00310; PTBI; 1.
KW Hypothetical protein.
SQ SEQUENCE 508 AA; 56794 MW; DPF8A818BFF8631 CRC64;

Query Match      98.4%; Score 685; DB 11; Length 508;
Best Local Similarity 97.7%; Pred. No. 7.3e-67;
Matches 127; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 DTVPDNRHNRKFKVINVDGNGELSGIMELTDTELLYTRKRDSSVKWHYLCRLRYGYDSN 60
DB 11 DTVPDNRHNRKFKVINVDGNGELSGIMELTDTELLYTRKRDSSVKWHYLCRLRYGYDSN 70
QY 61 LFSFESGRCCTGGGIFAFKCARABELFNMLQEIQNNSINNVVEPVPVVRNNHOTELEVP 120
DB 71 LFSFESGRCCTGGGIFAFKCARABELFNMLQEIQNNSINNVVEPVPVVRNNHOTELEVP 130
QY 121 RTPRTPTTPG 130
DB 131 RTPRTPTTPG 140

RESULT 3
Q90ZFS PRELIMINARY; PRT; 509 AA.
AC Q90ZFS;
DT 01-DEC-2001 (TREMELrel. 19, Created)
DT 01-DEC-2001 (TREMELrel. 19, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE FRS2.
GN XFRS2.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RA Kusakabe M., Masuyama N., Hanafusa H., Nishida E.;
RT "Xenopus FRS2 is involved in early embryogenesis in cooperation with the Src family kinase Lalco."
RL EMBO Reports 0:0-0(2001).
DR EMBL; AB064525; BAB61837.1; -
DR GO; GO:0005158; F:Insulin receptor binding; IEA.
DR InterPro; IPR002404; Insln_receptorS1.
DR Pfam; PF02174; IRS; 1.
DR SMART; SM00310; PTBI; 1.
SQ SEQUENCE 509 AA; 56963 MW; 9D9D56C276CD4041 CRC64;

Query Match      88.6%; Score 616.5; DB 13; Length 509;
Best Local Similarity 90.0%; Pred. No. 2.7e-59;
Matches 117; Conservative 7; Mismatches 5; Indels 1; Gaps 1;

QY 1 DTVPDNRHNRKFKVINVDGNGELSGIMELTDTELLYTRKRDSSVKWHYLCRLRYGYDSN 60
DB 11 ETIPDQNRKFKVINVDGNGELSGIMELTDTELLYTRKRDSSVKWHYLCRLRYGYDSN 70
QY 61 LFSFESGRCCTGGGIFAFKCARABELFNMLQEIQNNSINNVVEPVPVVRNNHOTELEVP 120
DB 71 LFSFESGRCCTGGGIFAFKCARABELFNMLQEIQNNSINNVVEPVPVVRNNHOTELEVP 129
QY 121 RTPRTPTTPG 130
DB 130 RTPRTPTTPG 139

RESULT 4
Q7ZWM2

```

```

ID Q7ZWM2 PRELIMINARY; PRT; 509 AA.
AC Q7ZWM2;
DT 01-JUN-2003 (Tremblrel. 24, Created)
DT 01-JUN-2003 (Tremblrel. 24, Last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE Similar to sucl-associated neurotrophic factor target (FGFR signalling adaptor).
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RA Klein S., Strausberg R.;
RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC049433; AAL4943.1; -.
DR GO; GO:0005158; F:insulin receptor binding; IEA.
DR InterPro; IPR002404; Insln_receptorS1.
DR Pfam; PF02174; IRS; 1.
DR SMART; SM00310; PTBI; 1.
DR SEQUENCE 509 AA; 56923 MW; 8F8E724A01CA3847 CRC64;

Query Match 88.6%; Score 616.5; DB 13; Length 509;
Best Local Similarity 90.0%; Pred. NO. 2.7e-59;
Matches 117; Conservative 7; Mismatches 5; Indels 1; Gaps 1;

QY 1 DTVPDNRHNFVKVINVDGNGELSGIMELTDTELLYTRKRDVSKWHYLCRLRYGYDSN 60
Db 11 ETIPDQNRFKVINVDGNGELSGIMELRENELLYTRKRDVSKWHYLCRLRYGYDSN 70
QY 61 LFSFESGRCQTGGIFAFKCARAEELFNNLQIMQNNSINVVEEPVVERNNHOTELEVP 120
Db 71 LFSFESGRCQTGGIFAFKCARAEELFNNLQIMQNNSINVVEEPVVERNNHOTELEVP 129

QY 121 RTPRTPTTPG 130
Db 130 RTPRTPTTPG 139

RESULT 5
Q8UVU3 PRELIMINARY; PRT; 509 AA.
AC Q8UVU3;
DT 01-MAR-2002 (Tremblrel. 20, Created)
DT 01-MAR-2002 (Tremblrel. 20, Last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE Sucl-associated neurotrophic factor target XSNT.
GN SNT.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RA Akagi K., Mood K., Daar I.O.;
RL "Fibroblast Growth Factor Receptor-Induced Mesoderm Formation in Xenopus Embryos Is Mediated by Adaptor Protein SNT.";
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF390895; AAL57304.1; -.
DR GO; GO:0005158; F:insulin receptor binding; IEA.
DR InterPro; IPR002404; Insln_receptorS1.
DR Pfam; PF02174; IRS; 1.
DR SMART; SM00310; PTBI; 1.
DR SEQUENCE 509 AA; 56965 MW; 7FB6CEA4DCD6DF5B CRC64;

Query Match 88.1%; Score 613.5; DB 13; Length 509;
Best Local Similarity 88.5%; Pred. NO. 5.7e-59;
Matches 115; Conservative 10; Mismatches 4; Indels 1; Gaps 1;

QY 1 DTVPDNRHNFVKVINVDGNGELSGIMELTDTELLYTRKRDVSKWHYLCRLRYGYDSN 60

```

```

Db 11 ETIPDQNRFKVINVDGNGELSGIMELTDTELLYTRKRDVSKWHYLCRLRYGYDSN 70
QY 61 LFSFESGRCQTGGIFAFKCARAEELFNNLQIMQNNSINVVEEPVVERNNHOTELEVP 120
Db 71 LFSFESGRCQTGGIFAFKCARAEELFNNLQIMQNNSINVVEEPVVERNNHOTELEVP 129

QY 121 RTPRTPTTPG 130
Db 130 RTPRTPTTPG 139

RESULT 6
Q43559 PRELIMINARY; PRT; 492 AA.
AC Q43559;
DT 01-JUN-1998 (Tremblrel. 06, Created)
DT 01-JUN-1998 (Tremblrel. 06, Last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE FGFR signalling adaptor SNT-2.
GN SNT-2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=98324992; PubMed=9660748;
RA Xu H., Lee K.W., Goldfarb M.;
RL "Novel recognition motif on fibroblast growth factor receptor mediates direct association and activation of SNT adaptor proteins.";
RL J. Biol. Chem. 273:17987-17990(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Strausberg R.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF036718; AAB92555.1; -.
DR EMBL; BC010611; AAH10611.1; -.
DR Genew; HGNC:16970; FR33.
DR GO; GO:0008543; P:FGF receptor signaling pathway; TAS.
DR GO; GO:0007165; P:signal transduction; TAS.
DR InterPro; IPR002404; Insln_receptorS1.
DR Pfam; PF02174; IRS; 1.
DR SMART; SM00310; PTBI; 1.
DR SEQUENCE 492 AA; 54461 MW; F06BFC62B531765 CRC64;

Query Match 73.3%; Score 510.5; DB 4; Length 492;
Best Local Similarity 70.2%; Pred. NO. 1.3e-47;
Matches 92; Conservative 21; Mismatches 17; Indels 1; Gaps 1;

QY 1 DTVPDNRHNFVKVINVDGNGELSGIMELTDTELLYTRKRDVSKWHYLCRLRYGYDSN 60
Db 11 DSVPDNHPFKVINVDGNGELSGVWELTQSELVJLHRRVAVRWYLCRLRYGYDSN 70
QY 61 LFSFESGRCQTGGIFAFKCARAEELFNNLQIMQNNSINVVEEPVVERNNHOTELEVP 119
Db 71 LFSFESGRCQTGGIFAFKCARAEELFNNLQIMQNNSINVVEEPVVERNNHOTELEVP 130

QY 120 RTPRTPTTPG 130
Db 131 PRAQPFNALG 141

RESULT 7
Q91WJ0 PRELIMINARY; PRT; 492 AA.
AC Q91WJ0;
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE Similar to sucl-associated neurotrophic factor target 2 (FGFR

```

DE signalling adaptor) (Fibroblast growth factor receptor substrate
 GN 3).
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Eye, and Retina;
 RA Strausberg R.;
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
 [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=129/SvJ;
 RA Zhou L., McDougall K., Kubu C.J., Verdi J.M., Meakin S.O.;
 RT "Genomic Organization and Comparative Sequence Analysis of the Mouse
 RT and Human FR32, FR33 genes."
 RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC014819; AAH14819.1; -
 DR EMBL; AF456480; AA015529.1; -
 DR MGD; MGI:2135965; FR33.
 DR GO; GO:0005158; F:insulin receptor binding; IEA.
 DR GO; GO:0004872; P:receptor activity; IEA.
 DR GO; GO:0006118; P:electron transport; IEA.
 DR InterPro; IPR001128; Cytochrome_P450.
 DR InterPro; IPR002404; Insln_receptorS1.
 DR Pfam; PF02174; IRS; 1.
 DR SMART; SM00310; PTB1; 1.
 DR PROSITE; PS00086; CYTOCHROME_P450; 1.
 DR Receptor.
 KW Receptor.
 SQ SEQUENCE 492 AA; 53976 MW; COA895B9173394B6 CRC64;
 Query Match 69.6%; Score 484.5; DB 11; Length 492;
 Best Local Similarity 69.3%; Pred. No. 9.3e-45;
 Matches 88; Conservative 21; Mismatches 17; Indels 1; Gaps 1;
 QY 1 DTVPDHRNFKVINDDGNGELSGIMELTDTELILYTRKDSVKWVHLCRLRYGYDSN 60
 Db 11 DSVPHNPTKFKVTNDDGVELSGVMELTQSELVHLHOREAVRWPLYCLRLRYGYDSN 70
 QY 61 LFSFSGRRCTGGGIFAFKCAAEELFNLMQNMNSINVTPEPV-VERNNHQTLEV 119
 Db 71 LFSFSGRRCTGGGIFAFKCAAEELFNLMQNMNSINVTPEPV-VERNNHQTLEV 130
 QY 120 PRTPPT 126
 Db 131 PRGPPQ 137
 RESULT 8
 ID O44830 PRELIMINARY; PRT; 195 AA.
 AC O44830;
 DT 01-JUN-1998 (TrEMBLrel. 06, Created)
 DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE F54D12.6 protein.
 GN F54D12.6.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Peioderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RX MEDLINE=94150718; PubMed=7906398;
 RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
 RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
 RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
 RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnson L.,
 RA Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,
 RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,

RA Parsons J., Percy C., Rifkin L., Roopra A., Saunders D., Showkeen R.,
 RA Smalton N., Smith A., Sonhammer E., Staden R., Sulston J.,
 RA Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
 RA Watson A., Weinstock L., Wilkinson-Sproat J., Wohlman P.,
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
 RT elegans."
 RL Nature 368:32-38 (1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RA Maggi L., Goela D.;
 RT "The sequence of C. elegans cosmid F54D12."
 RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
 [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RA Waterston R.;
 RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF040647; AAB94991.1; -
 DR PIR; T32819; T32819
 DR WormPep; F54D12.6; CE17871.
 DR GO; GO:0005158; F:insulin receptor binding; IEA.
 DR InterPro; IPR002404; Insln_receptorS1.
 DR InterPro; IPR004008; Reg_chr_condens.
 DR Pfam; PF02174; IRS; 1.
 DR SMART; SM00310; PTB1; 1.
 DR PROSITE; PS00526; RC1.2; 1.
 DR PROSITE; PS00526; RC1.2; 1.
 SQ SEQUENCE 195 AA; 22009 MW; 7E477ABC2B3695D7 CRC64;
 Query Match 25.6%; Score 178; DB 5; Length 195;
 Best Local Similarity 32.5%; Pred. No. 1.9e-11;
 Matches 39; Conservative 30; Mismatches 41; Indels 10; Gaps 3;
 QY 7 HENKFKVINDDDGNGELSGIMELTDTELILYTRKDSVKWVHLCRLRYGYDS-NLFSFE 65
 Db 42 HGNKK-----NELVHAWLRVTRTTITLISKCECLVWPLPLIRRYGTSAGIFFE 93
 QY 66 SGRRCTGGGIFAFKCAAEELFNLMQNMNSINVTPEPV-VERNNHQTLEV 124
 Db 94 SGRRCTGGGIFAFKCAAEELFNLMQNMNSINVTPEPV-VERNNHQTLEV 153
 RESULT 9
 ID Q9VLL4 PRELIMINARY; PRT; 442 AA.
 AC Q9VLL4;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE CG13398 protein.
 GN CG13398.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Berkley;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celnik S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
 RA Wan X.H., Doyle C., Baxter E.G., Helt J., Nelson C.R., Miklos G.L.G.,
 RA Abail J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.V., Benos P.V., Bernan B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,

```
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foslter C., Gabrielle A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Idegam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirkas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-f., Wasserman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RA "The genome sequence of Drosophila melanogaster.";
RA Science 287:2185-2195(2000).
RL EMBL; AB003621; AAF2573.1; -.
DR FlyBase; FBgn0032042; CG13398.
DR GO; GO:0005158; F:insulin receptor binding; IEA.
DR InterPro; IPR002404; Insnl_receptorS1.
DR Pfam; PF02174; IRS; 1.
DR SMART; SM00310; PTB1; 1.
DR SQU SEQUENCE 442 AA; 48240 MW; 95853139E7688D69 CRC64;

Query Match 25.0%; Score 174; DB 5; Length 442;
Best Local Similarity 34.4%; Pred. No. 1.5e-10;
Matches 42; Conservative 22; Mismatches 52; Indels 6; Gaps 2;

QY 7 HRNKFV-----INVDDGNGELGSGIMELTDTLLILYTRKDSVKVHYLCRLRYGYDSNLF 62
DB 16 HENFVRVVAHLQPTGPTPIIRSGYLELTRELIFQSPGCEPIVWALQHLRRYGLNDLF 75

QY 63 SFESGRRCQTGGIFAFKARAEELFNMQLQIMQNSINVVEPVRN--HOTELEVP 120
DB 76 SFEAGRCMSGPGIYTPRVHNAEQLYPMFQRYINAVNTDAFQGERERVNSAHSVSNMG 135

QY 121 RT 122
DB 136 RT 137

RESULT 10
Q8MT60 Q8MT60 PRELIMINARY; PRT; 442 AA.
AC Q8MT60;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE RE08678p.
GN Drosophila melanogaster (Fruit fly).
OS Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
CX NCBI_TaxID=7227;
EN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkely;
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champagne M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
RA George R., Gonzalez M., Guarin H., Kronmiller B., Li P., Liao G.,
RA Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragas V., Park S.,
RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA Ceiniker S.;

Query Match 25.0%; Score 174; DB 5; Length 442;
Best Local Similarity 34.4%; Pred. No. 1.5e-10;
Matches 42; Conservative 22; Mismatches 52; Indels 6; Gaps 2;

QY 7 HRNKFV-----INVDDGNGELGSGIMELTDTLLILYTRKDSVKVHYLCRLRYGYDSNLF 62
DB 16 HENFVRVVAHLQPTGPTPIIRSGYLELTRELIFQSPGCEPIVWALQHLRRYGLNDLF 75

QY 63 SFESGRRCQTGGIFAFKARAEELFNMQLQIMQNSINVVEPVRN--HOTELEVP 120
DB 76 SFEAGRCMSGPGIYTPRVHNAEQLYPMFQRYINAVNTDAFQGERERVNSAHSVSNMG 135

QY 121 RT 122
DB 136 RT 137

RESULT 10
Q8MT60 Q8MT60 PRELIMINARY; PRT; 442 AA.
AC Q8MT60;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE RE08678p.
GN Drosophila melanogaster (Fruit fly).
OS Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
CX NCBI_TaxID=7227;
EN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkely;
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champagne M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
RA George R., Gonzalez M., Guarin H., Kronmiller B., Li P., Liao G.,
RA Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragas V., Park S.,
RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA Ceiniker S.;
```

```
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY118363; AAU48392.1; -.
DR FlyBase; FBgn0032042; CG13398.
DR GO; GO:0005158; F:insulin receptor binding; IEA.
DR InterPro; IPR002404; Insnl_receptorS1.
DR Pfam; PF02174; IRS; 1.
DR SMART; SM00310; PTB1; 1.
DR SQU SEQUENCE 442 AA; 48217 MW; FD67D8BE3A5BF038 CRC64;

Query Match 25.0%; Score 174; DB 5; Length 442;
Best Local Similarity 34.4%; Pred. No. 1.5e-10;
Matches 42; Conservative 22; Mismatches 52; Indels 6; Gaps 2;

QY 7 HRNKFV-----INVDDGNGELGSGIMELTDTLLILYTRKDSVKVHYLCRLRYGYDSNLF 62
DB 16 HENFVRVVAHLQPTGPTPIIRSGYLELTRELIFQSPGCEPIVWALQHLRRYGLNDLF 75

QY 63 SFESGRRCQTGGIFAFKARAEELFNMQLQIMQNSINVVEPVRN--HOTELEVP 120
DB 76 SFEAGRCMSGPGIYTPRVHNAEQLYPMFQRYINAVNTDAFQGERERVNSAHSVSNMG 135

QY 121 RT 122
DB 136 RT 137

RESULT 11
Q9QZK7 Q9QZK7 PRELIMINARY; PRT; 444 AA.
AC Q9QZK7;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Adaptor protein (Inhibitory adapter molecule DOK3).
GN DOKL OR DOK3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALE/C;
RX MEDLINE=20036788; PubMed=10567556;
RA Cong F., Yuan B., Goff S.P.;
RA "Characterization of a Novel Member of the DOK Family that Binds and
RT Modulates Abl Signaling.";
RL Mol. Cell. Biol. 19:8314-8325(1999).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=20198458; PubMed=10733577;
RA Lemay S., Davidson D., Latour S., Veillette A.;
RA "Dok-3, a novel adapter molecule involved in the negative regulation
RT of immunoreceptor signaling.";
RL Mol. Cell. Biol. 20:2743-2754(2000).
DR EMBL; AF179242; AAF14285.1; -.
DR EMBL; AF237580; AAF61309.1; -.
DR MGD; MGI:1351490; Dok1.
DR GO; GO:0005737; C:cytoplasm; IDA.
DR GO; GO:0007265; P:RAS protein signal transduction; IDA.
DR InterPro; IPR002404; F:protein binding; IPI.
DR InterPro; IPR001849; PH.
DR Pfam; PF02174; IRS; 1.
DR Pfam; PF00169; PH; 1.
DR SMART; SM00233; PH; 1.
DR SMART; SM00310; PTB1; 1.
DR SQU SEQUENCE 444 AA; 48027 MW; 124E705712202D9B CRC64;

Query Match 20.3%; Score 141; DB 11; Length 444;
Best Local Similarity 28.3%; Pred. No. 6.5e-07;
Matches 36; Conservative 20; Mismatches 47; Indels 24; Gaps 3;

QY 23 LGSIMELTDTLLILYTRKDSVKVHYLCRLRYGYDSNLFSESGRCQTGGIFAFKCA 82
```

```
Db 184 LQDDIQRET-----SPQACFSFYRFLRKYSGDKGVFSFAGRCDCSGEGLFAFSSP 238
QY 93 RAEEELNMLQEIOMQNSINVVE-----EPVVERNNHQTELEVPRTPR 124
Db 239 RAPDTCGVAAIAEQRERLELAMSPPCPLPRALSPLSLEPPGELREVARGFELPTPRK 298
QY 125 TP-TTPG 130
Db 299 LPLTDFG 305

RESULT 12
Q8WU28 PRELIMINARY; PRT; 223 AA.
AC Q8WU28;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Eye;
RA Strausberg R.;
RL Submitted (DSC-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC019045; AAH19045.1; -.
DR GO; GO:0005158; F:insulin receptor binding; IEA.
DR InterPro; IPR002404; Insln_receptorS1.
DR Pfam; PF02174; IRS; 1.
DR SMART; SM00310; PTBI; 1.
KW Hypothetical protein.
SQ SEQUENCE 223 AA; 25538 MW; DFD6F27B1981AEBB CRC64;

Query Match 19.8%; Score 137.5; DB 4; Length 223;
Best Local Similarity 31.5%; Pred. No. 6.8e-07;
Matches 34; Conservative 19; Mismatches 42; Indels 13; Gaps 3;

QY 24 GSGIMELTDTELILYTRKRSVK---WHYLCRLRYGYDSNLFSPESGRCRCQTGGIFAPK 80
Db 45 GECTMQITHENIYLDIHNKVLVWPLSSLRRYGRDSTWFTFESGRMCDTGEGLFTFQ 104
QY 81 CARAEELFNMLQEIOMQNSINVVE-----PVERNNHQTELEVPR 122
Db 105 TREGE---MIYQKVHSATLAIAEQHERLMEQKARLQTSLTEPMT 148

RESULT 13
Q96LU2 PRELIMINARY; PRT; 223 AA.
AC Q96LU2;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein FLJ25066.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Cerebellum;
RA Nishi T., Nakagawa S., Senoh A., Mizuguchi H., Inagaki H., Suzuki Y.,
RA Hata H., Nakagawa K., Mizuno S., Morinaga M., Kawamura M.,
RA Sugiyama T., Irie R., Otsuki T., Sato H., Nishikawa T., Sugiyama A.,
RA Kawakami B., Nagai K., Isogai T., Sugano S.;
RT "NEDO human cDNA sequencing project."
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK057795; BAB7157.1; -.
DR GO; GO:0005158; F:insulin receptor binding; IEA.
```

```
DR InterPro; IPR002404; Insln_receptorS1.
DR Pfam; PF02174; IRS; 1.
DR SMART; SM00310; PTBI; 1.
KW Hypothetical protein.
SQ SEQUENCE 223 AA; 25506 MW; 2B87B77A499B05A1 CRC64;

Query Match 19.8%; Score 137.5; DB 4; Length 223;
Best Local Similarity 31.5%; Pred. No. 6.8e-07;
Matches 34; Conservative 19; Mismatches 42; Indels 13; Gaps 3;

QY 24 GSGIMELTDTELILYTRKRSVK---WHYLCRLRYGYDSNLFSPESGRCRCQTGGIFAPK 80
Db 45 GECTMQITHENIYLDIHNKVLVWPLSSLRRYGRDSTWFTFESGRMCDTGEGLFTFQ 104
QY 81 CARAEELFNMLQEIOMQNSINVVE-----PVERNNHQTELEVPR 122
Db 105 TREGE---MIYQKVHSATLAIAEQHERLMEQKARLQTSLTEPMT 148

RESULT 14
Q96HI2 PRELIMINARY; PRT; 257 AA.
AC Q96HI2;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Strausberg R.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC008583; AAH08583.1; -.
DR GO; GO:0005158; F:insulin receptor binding; IEA.
DR InterPro; IPR002404; Insln_receptorS1.
DR Pfam; PF02174; IRS; 1.
DR SMART; SM00310; PTBI; 1.
KW Hypothetical protein.
FT NON_TER 1
SQ SEQUENCE 257 AA; 29455 MW; 931D34A8F12A952A CRC64;

Query Match 19.8%; Score 137.5; DB 4; Length 257;
Best Local Similarity 31.5%; Pred. No. 8.1e-07;
Matches 34; Conservative 19; Mismatches 42; Indels 13; Gaps 3;

QY 24 GSGIMELTDTELILYTRKRSVK---WHYLCRLRYGYDSNLFSPESGRCRCQTGGIFAPK 80
Db 79 GECTMQITHENIYLDIHNKVLVWPLSSLRRYGRDSTWFTFESGRMCDTGEGLFTFQ 138
QY 81 CARAEELFNMLQEIOMQNSINVVE-----PVERNNHQTELEVPR 122
Db 139 TREGE---MIYQKVHSATLAIAEQHERLMEQKARLQTSLTEPMT 182

RESULT 15
Q8NSA4 PRELIMINARY; PRT; 412 AA.
AC Q8NSA4;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Docking protein 2, 56kD.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
```

RA Strausberg R.;
 RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
 CC -!- SIMILARITY: CONTAINS 1 PH DOMAIN.
 DR EMBL; BC032623; AAH32623.1; -.
 DR GO; GO:0005158; F:insulin receptor binding; IEA.
 DR InterPro; IPR002404; Insln_receptorS1.
 DR InterPro; IPR001849; PH.
 DR Pfam; PF02174; IRS; 1.
 DR Pfam; PF00169; PH; 1.
 DR SMART; SM00233; PH; 1.
 DR SMART; SM00310; PTBI; 1.
 DR PROSITE; PS50003; PH DOMAIN; 1.
 SQ SEQUENCE 412 AA; 35378 MW; EFAEDBA68439757F CRC64;

Query Match 19.4%; Score 135; DB 4; Length 412;
 Best Local Similarity 34.0%; Pred. No. 2.7e-06;
 Matches 33; Conservative 13; Mismatches 33; Indels 18; Gaps 3;

QY	47	WHYLCRLRYGYDSNLFESGRRRCOTGGIGFAPKCARAEELFNMLQEIIM--QNNINVVE	104
Db	194	WFYRFLRRFGDKVTSFEAGRCVSGEGNFEFETRQGNIEIFLALEEAIQAQNAAPATP	253
QY	105	EPVVERNHHQTELEVR-----TPRTPTTP	129
Db	254	QP----QPATIPASLRPRDPSYSRPHDSLPPSPITP	286

Search completed: August 2, 2004, 09:31:15
 Job time : 60.1579 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 2, 2004, 09:27:44 ; Search time 24.8026 Seconds
(without alignments)
270.591 Million cell updates/sec

Title: US-09-757-415a-1_COPY_11_140
Perfect score: 696
Sequence: 1 DTPVDNRHNRKFKVNVDDG.....NNHQTELEVPRTPTPTTPG 130

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: /cgn2_6/prodata/2/iaa/5A_COMB.pep.*
2: /cgn2_6/prodata/2/iaa/5B_COMB.pep.*
3: /cgn2_6/prodata/2/iaa/6A_COMB.pep.*
4: /cgn2_6/prodata/2/iaa/6B_COMB.pep.*
5: /cgn2_6/prodata/2/iaa/PCTUS_COMB.pep.*
6: /cgn2_6/prodata/2/iaa/backfiles1.pep.*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	685	98.4	508	4	US-08-980-523-9
2	679	97.6	129	4	US-08-980-523-11
3	143.5	20.6	481	3	US-08-787-091-2
4	78	11.2	1234	2	US-08-317-310A-15
5	78	11.2	1234	5	PCT-US95-13041-15
6	78	11.2	1242	4	US-09-508-691-1
7	78	11.2	1243	2	US-08-557-139-2
8	77.5	11.1	112	4	US-08-980-523-10
9	77	11.1	38	3	US-08-787-091-9
10	75	10.8	435	2	US-08-531-439B-4
11	74.5	10.7	709	4	US-09-489-847-132
12	72.5	10.4	541	4	US-09-134-001C-4481
13	68.5	9.8	342	4	US-09-543-681A-5131
14	68.5	9.8	1561	3	US-08-994-017-23
15	68.5	9.8	1561	4	US-09-456-474-23
16	68	9.8	159	4	US-09-508-691-5
17	68	9.8	1321	2	US-08-317-310A-64
18	67.5	9.7	118	4	US-08-936-165A-464
19	67.5	9.7	342	4	US-09-489-039A-14092
20	67	9.6	144	4	US-09-134-000C-4435
21	67	9.6	413	4	US-09-215-694-6
22	67	9.6	492	4	US-09-134-001C-3895
23	66.5	9.6	1817	4	US-09-004-838-125
24	66	9.5	378	4	US-09-134-001C-5487
25	65.5	9.4	962	4	US-09-134-001C-4497
26	65	9.3	283	4	US-09-198-452A-1116
27	65	9.3	1890	4	US-09-004-838-88

28 65 9.3 2807 4 US-09-543-681A-4980
29 64.5 9.3 290 4 US-09-910-174B-8
30 64.5 9.3 290 4 US-09-620-461-8
31 64 9.2 231 4 US-09-543-681A-7920
32 64 9.2 242 2 US-08-553-497A-26
33 63.5 9.1 86 2 US-08-858-767-35
34 63.5 9.1 86 2 US-08-863-028-35
35 63.5 9.1 291 2 US-08-531-927B-4
36 63.5 9.1 360 2 US-08-531-927B-2
37 63.5 9.1 360 3 US-09-041-886-13
38 63.5 9.1 527 3 US-08-369-822C-25
39 63.5 9.1 527 3 US-08-582-776C-40
40 63.5 9.1 527 3 US-08-434-831B-37
41 63.5 9.1 730 3 US-09-398-865A-2
42 63.5 9.1 730 4 US-09-710-714-2
43 63 9.1 87 4 US-09-107-532A-5948
44 62.5 9.0 252 4 US-09-134-000C-5573
45 62.5 9.0 913 4 US-09-623-624-2

ALIGNMENTS

RESULT 1
US-08-980-523-9
; Sequence 9, Application US/08980523
; Patent No. 6310181
; GENERAL INFORMATION:
; APPLICANT: Kouhara, Haruhiko
; APPLICANT: Spivak-Kroizman, Taly
; APPLICANT: Iax, Irit
; APPLICANT: Schlössinger, Joseph
; TITLE OF INVENTION: ADAPTOR PROTEIN FRG2 AND
; TITLE OF INVENTION: RELATED PRODUCTS AND METHODS
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; STREET: Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
MEDIUM TYPE: storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: FASTSEQ for Windows 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/980.523
FILING DATE: December 1, 1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US97/21851
FILING DATE: December 1, 1997
APPLICATION NUMBER: 60/032,093
FILING DATE: December 3, 1996
ATTORNEY/AGENT INFORMATION:
NAME: Warburg, Richard J.
REGISTRATION NUMBER: 32,327
REFERENCE/DOCKET NUMBER: 230/045
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 508 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein

Sequence 4980, Ap
Sequence 8, Appli
Sequence 8, Appli
Sequence 7820, Ap
Sequence 26, Appl
Sequence 35, Appl
Sequence 35, Appl
Sequence 4, Appli
Sequence 2, Appli
Sequence 13, Appl
Sequence 25, Appl
Sequence 40, Appl
Sequence 37, Appl
Sequence 2, Appli
Sequence 5948, Ap
Sequence 5573, Ap
Sequence 2, Appli

US-08-980-523-9

Query Match 98.4%; Score 685; DB 4; Length 508;
Best Local Similarity 97.7%; Pred. No. 4.4e-77;
Matches 127; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 DTVPDNRKFKVINVDGNGELSGIMELTDTELLIYTRKDSVKWHYLCRRYGYDSN 60
DB 11 DTVPDNRKFKVINVDGNGELSGIMELTDTELLIYTRKDSVKWHYLCRRYGYDSN 70
QY 61 LFSFSGRRCTGGGIFAFKCARAEELFNMLOEIMONNSINVVEPVRNNHOTELEVP 120
DB 71 LFSFSGRRCTGGGIFAFKCARAEELFNMLOEIMONNSINVVEPVRNNHOTELEVP 130
QY 121 RTRPTPTTPG 130
DB 131 RTRPTPTTPG 140

RESULT 2

US-08-980-523-11
; Sequence 11, Application US/08980523
; Patent No. 6310181

GENERAL INFORMATION:

APPLICANT: Kouhara, Haruhiko
APPLICANT: Spivak-Kroizman, Taly
APPLICANT: Lax, Irit
APPLICANT: Schliesinger, Joseph
TITLE OF INVENTION: ADAPTOR PROTEIN FR52 AND
TITLE OF INVENTION: RELATED PRODUCTS AND METHODS
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
STREET: Suite 4700
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071-2066

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
MEDIUM TYPE: storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: FastSO for Windows 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/980,523
FILING DATE: December 1, 1997

CLASSIFICATION: 435

PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US97/21851

FILING DATE: December 1, 1997

APPLICATION NUMBER: 60/032,093

FILING DATE: December 3, 1996

ATTORNEY/AGENT INFORMATION:

NAME: Warburg, Richard J.

REGISTRATION NUMBER: 32,327

REFERENCE/DOCKET NUMBER: 230/045

TELECOMMUNICATION INFORMATION:

TELEPHONE: (213) 489-1600

TELEFAX: (213) 955-0440

TELEX: 67-3510

INFORMATION FOR SEQ ID NO: 11:

SEQUENCE CHARACTERISTICS:

LENGTH: 129 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-980-523-11

Query Match 97.6%; Score 679; DB 4; Length 129;
Best Local Similarity 97.7%; Pred. No. 3.6e-77;

Matches 126; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 DTVPDNRKFKVINVDGNGELSGIMELTDTELLIYTRKDSVKWHYLCRRYGYDSN 60
DB 1 DTVPDNRKFKVINVDGNGELSGIMELTDTELLIYTRKDSVKWHYLCRRYGYDSN 60
QY 61 LFSFSGRRCTGGGIFAFKCARAEELFNMLOEIMONNSINVVEPVRNNHOTELEVP 120
DB 61 LFSFSGRRCTGGGIFAFKCARAEELFNMLOEIMONNSINVVEPVRNNHOTELEVP 120
QY 121 RTRPTPTTP 129
DB 121 RTRPTPTTP 129

RESULT 3

US-08-787-091-2

; Sequence 2, Application US/08787091

; Patent No. 6100386

GENERAL INFORMATION:

APPLICANT: Carpino, Nicholas A.
APPLICANT: Kobayashi, Ryuji
APPLICANT: Wisniewski, David G.
APPLICANT: Strife, Annabel O'C.
APPLICANT: Clarkson, Bayard D.

TITLE OF INVENTION: A No. 6100386el Human Gene/Protein Involved in

TITLE OF INVENTION: Chronic Myelogenous Leukemia

NUMBER OF SEQUENCES: 20

CORRESPONDENCE ADDRESS:

ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.

STREET: Two Militia Drive

CITY: Lexington

STATE: MA

COUNTRY: US

ZIP: 02173

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/787,091

FILING DATE:

CLASSIFICATION: 530

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 60/030,418

FILING DATE: 01-NOV-1996

ATTORNEY/AGENT INFORMATION:

NAME: Granahan, Patricia

REGISTRATION NUMBER: 32,227

REFERENCE/DOCKET NUMBER: CSHL96-05pa

TELECOMMUNICATION INFORMATION:

TELEPHONE: 617-861-6240

TELEFAX: 617-861-9540

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 481 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-787-091-2

Query Match 20.6%; Score 143.5; DB 3; Length 481;

Best Local Similarity 23.5%; Pred. No. 2.6e-09;

Matches 33; Conservative 20; Mismatches 50; Indels 9; Gaps 2;

QY 24 GSGIMELTDTELLIYTRKDS-----VKWHYLCRRYGYDSNLFSPESGRRCTGGGIF 77
DB 172 GSYLVIRVEAERLTLLTVGAOSQILEPLSLWPTLLRRYGRDKVMVFSFAGRRCPSPGGTF 231
QY 78 AFKCARAEELFNMLOEIMONNSINVVE---EPVVRNNHOTELEVPRTPTP 126
DB 232 TFOAQNDIFQAVETAIRHQAGKAGQGHVLRADSHGEVAEGKLPSP 283

```

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/13041
FILING DATE: Herewith
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/317,310
FILING DATE: 03-OCT-1994
ATTORNEY/AGENT INFORMATION:
NAME: Louis Myers
REGISTRATION NUMBER: 35,965
REFERENCE/DOCKET NUMBER: JDP-022PC
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 1234 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
FRAGMENT TYPE: internal
PCT-US95-13041-15

Query Match 11.2%; Score 78; DB 5; Length 1234;
Best Local Similarity 30.0%; Pred. No. 1.6;
Matches 21; Conservative 20; Indels 22; Gaps 1;

Qy 52 LRRYGVDSNLFSPESGRRCOTGGCIFAFCARAEELFNLQEIIMQNNSINNVVEEPVVERN 111
Db 206 IRRCHSENFFFEVGRSAVTGGEFFWQ-----VDDSVVAQN 243
Qy 112 NHOELEVP 121
Db 244 MHEITLEAMR 253

RESULT 6
US-09-508-691-1
Sequence 1, Application US/09508691
Patent No. 6498139
GENERAL INFORMATION:
APPLICANT: YAZAKI, YOSHIO
APPLICANT: ASANO, TOMOICHIRO
APPLICANT: KUBO, HIDEO
APPLICANT: KANDA, AKIRA
TITLE OF INVENTION: REMEDIES FOR DISEASES CAUSED BY INSULIN RESISTANCE
FILE REFERENCE: 4895-0019-OPCT
CURRENT APPLICATION NUMBER: US/09/508,691
CURRENT FILING DATE: 2000-03-29
PRIOR APPLICATION NUMBER: PCT/JF98/04293
PRIOR FILING DATE: 1998-09-25
PRIOR APPLICATION NUMBER: JF9-263719
PRIOR FILING DATE: 1997-09-29
NUMBER OF SEQ ID NOS: 5
SOFTWARE: PatentIn version 3.0
SEQ ID NO 1
LENGTH: 1242
TYPE: PRT
ORGANISM: Homo sapiens
US-09-508-691-1

Query Match 11.2%; Score 78; DB 4; Length 1242;
Best Local Similarity 30.0%; Pred. No. 1.6;
Matches 21; Conservative 7; Mismatches 20; Indels 22; Gaps 1;

Qy 52 LRRYGVDSNLFSPESGRRCOTGGCIFAFCARAEELFNLQEIIMQNNSINNVVEEPVVERN 111
Db 211 IRRCHSENFFFEVGRSAVTGGEFFWQ-----VDDSVVAQN 248

```

```
QY 112 NHQTELEVR 121
DB 249 MHETILEAMR 258

RESULT 7
US-08-557-139-2
; Sequence 2, Application US/08557139
; Patent No. 5827730
; GENERAL INFORMATION:
; APPLICANT: Pedersen, Oluf
; APPLICANT: Bjorbak, Christian
; APPLICANT: Frederiksen, Kathrine A.
; TITLE OF INVENTION: MUTANT DNA ENCODING INSULIN RECEPTOR
; TITLE OF INVENTION: SUBSTRATE 1
; NUMBER OF SEQUENCES: 41
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 58277300 No. 5827730disk of No. 5827730th America
; STREET: 405 Lexington Avenue
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10174
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/557,139
; FILING DATE: 12-FEB-1996
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Lambiris, Elias J.
; REGISTRATION NUMBER: 33,728
; REFERENCE/DOCKET NUMBER: 4041.204-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 867-0123
; TELEFAX: (212) 878-9655
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1243 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-557-139-2

Query Match 11.28; Score 78; DB 2; Length 1243;
Best Local Similarity 30.0%; Pred. No. 1.6;
Matches 21; Conservative 7; Mismatches 20; Indels 22; Gaps 1;

QY 52 LRRYGYDNLFSFSGRRCTGGTGFAPKCARAEELFNMLQELMKNNSINNVVEPVPVERN 111
DB 212 IRRCGHSENFIEVGRSAVTGPGFWMQ-----VDDSVVAQN 249

QY 112 NHQTELEVR 121
DB 250 MHETILEAMR 259

RESULT 8
US-08-980-523-10
; Sequence 10, Application US/08980523
; Patent No. 6310181
; GENERAL INFORMATION:
; APPLICANT: Kouhara, Haruhiko
; APPLICANT: Spivak-Kroizman, Taly
; APPLICANT: Lax, Irit
; APPLICANT: Schlessinger, Joseph
; TITLE OF INVENTION: ADAPTOR PROTEIN FR52 AND
; TITLE OF INVENTION: RELATED PRODUCTS AND METHODS
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:

QY 112 NHQTELEVR 121
DB 95 MHETILEAMR 104

Query Match 11.11; Score 77.5; DB 4; Length 112;
Best Local Similarity 30.0%; Pred. No. 0.063;
Matches 21; Conservative 7; Mismatches 19; Indels 23; Gaps 1;

QY 52 LRRYGYDNLFSFSGRRCTGGTGFAPKCARAEELFNMLQELMKNNSINNVVEPVPVERN 111
DB 58 IRRCGHSENFIEVGRSAVTGPGFEW-----MVDDSVVAQN 94

RESULT 9
US-08-787-091-9
; Sequence 9, Application US/08787091
; Patent No. 6100386
; GENERAL INFORMATION:
; APPLICANT: Carpino, Nicholas A.
; APPLICANT: Kobayashi, Ryuji
; APPLICANT: Wisniewski, David G.
; APPLICANT: Strife, Annabel O.C.
; APPLICANT: Clarkson, Bayard D.
; TITLE OF INVENTION: A No. 6100386el Human Gene/Protein Involved in
; TITLE OF INVENTION: Chronic Myelogenous Leukemia
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: MA
; COUNTRY: US
; ZIP: 02173
; COMPUTER READABLE FORM:
```

```

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/787,091
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/030,418
; FILING DATE: 01-NOV-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Granahan, Patricia
; REGISTRATION NUMBER: 32,227
; REFERENCE/DOCKET NUMBER: CSH196-059a
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-861-6240
; TELEFAX: 617-861-9540
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 38 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide
; US-08-787-091-9

Query Match 11.8; Score 77; DB 3; Length 38;
Best Local Similarity 40.6%; Pred. No. 0.016;
Matches 13; Conservative 9; Mismatches 10; Indels 0; Gaps 0;

QY 61 LFSFGRRRCQTGGIGFAFKCARAEELFNMLQ 92
Db 1 MFSFEAGRRXPSGGITFTQTAGNDIFQAVE 32

RESULT 10
US-08-531-439B-4
; Sequence 4, Application US/08531439B
; Patent No. 5981702
; GENERAL INFORMATION:
; APPLICANT: Zhang, Hui
; APPLICANT: Beach, David
; TITLE OF INVENTION: Cyclin/CDK Associated Proteins,
; TITLE OF INVENTION: and Uses Related Thereto
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley, Hoag & Eliot
; STREET: One Post Office Square
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII (text)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/531,439B
; FILING DATE: 21-SEP-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: CSV-006.01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 832-1000
; TELEFAX: (617) 832-7000
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 435 amino acids
; TYPE: amino acid
;
; MEDIUM TYPE: linear
; MOLECULE TYPE: protein
; US-08-531-439B-4

Query Match 10.8; Score 75; DB 2; Length 435;
Best Local Similarity 23.8%; Pred. No. 0.87;
Matches 40; Conservative 22; Mismatches 59; Indels 48; Gaps 8;

QY 4 PDNHNKFKVINDDGNGELSGIME-LTDTELI-----LYTRKEDSVKHYVLCRR 54
Db 97 PALNRENFPGVSDSLPDELLGLGFSCLCLELLKVSVCRWYRLASDESJWQTLDT- 155
QY 55 YGYDSNLFSPESGRRRCQTGGIGFAFKCARA-----BELFNMIOEI--- 94
Db 156 ---GKNLHPDVTGR--LLSQGVIAFRCPSPMDQPLAEHFSPFRVQMDLSNVIEVSTL 210
QY 95 -----MNNNSINV-VEEPVVE---RNHHQTELEVPRTPRTPTTP 129
Db 211 HGILSQCSKLQNLSELESLSDPIVNTIAKNSNLVRLNLPFCGCPKFP 258

RESULT 11
US-09-489-847-132
; Sequence 132, Application US/09489847
; Patent No. 6476195
; GENERAL INFORMATION:
; APPLICANT: Rosen et al
; TITLE OF INVENTION: 98 Human Secreted Proteins
; FILE REFERENCE: PZ031P1
; CURRENT APPLICATION NUMBER: US/09/489,847
; CURRENT FILING DATE: 2000-01-24
; EARLIER APPLICATION NUMBER: PCT/US99/17130
; EARLIER FILING DATE: 1999-07-29
; EARLIER APPLICATION NUMBER: 60/094,657
; EARLIER FILING DATE: 1998-07-30
; EARLIER APPLICATION NUMBER: 60/095,486
; EARLIER FILING DATE: 1998-08-05
; EARLIER APPLICATION NUMBER: 60/096,319
; EARLIER FILING DATE: 1998-08-12
; EARLIER APPLICATION NUMBER: 60/095,454
; EARLIER FILING DATE: 1998-08-06
; EARLIER APPLICATION NUMBER: 60/095,455
; EARLIER FILING DATE: 1998-08-06
; NUMBER OF SEQ ID NOS: 376
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 132
; LENGTH: 709
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (189)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (275)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (414)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (438)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (641)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (643)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

```

```
;
; NAME/KEY: SITE
; LOCATION: (696)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
;
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (697)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-489-847-132

Query Match
Best Local Similarity 10.7%; Score 74.5; DB 4; Length 709;
Matches 30; Conservative 17; Mismatches 53; Indels 19; Gaps 4;

QY 26 GIMELT-DTELILYTRKRD-----SVKWHYLCRLRRYGYDSNLFSPESGRRCTQGQ 75
DB 545 GVFTVPTDKSVYLRTNWRGULPSLTSVSNISVPRDQVACLTFFKERSGVVVCQIGRA 604
QY 76 --IPAFKCARAEELFNMQLQIMONNS-----INVVEPVPVVRNNHQTLEVPRTPT 125
DB 605 FMIIQQRTRAEIIFSLDEDVLPKPSFHHHSFWVNIIXNSPTSGKQLDLLFVTLTPT 663

RESULT 12
US-09-134-001C-4481
; Sequence 4481, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134.001C
; PRIOR FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 4481
; LENGTH: 541
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-4481

Query Match
Best Local Similarity 10.4%; Score 72.5; DB 4; Length 541;
Matches 31; Conservative 21; Mismatches 37; Indels 31; Gaps 7;

QY 10 KFKVINDDGNEGLSGIMELTDTELILYTRKRDVSKWHYLCRLRRYGYDSNLFSPESGRR 69
DB 26 KFRVLQV--GGHDLGS-----LFNQK-SNVEWDY-----FDVGLDFPESGYQ 64
QY 70 -----CQTGGQIFAPFKCAR-ABEELFNMQLQIMONNSINVVEPVPVVRNNHQTLEVP 121
DB 65 DVVVHILDEKQDFVFVREKYSLSMKLLS--LVSPYNTVIDNEYWDNQYQDQKTQIR 122

RESULT 13
US-09-543-681A-5131
; Sequence 5131, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543.681A
; PRIOR FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 5131
```

```
;
; LENGTH: 342
; TYPE: PRT
; ORGANISM: Proteus mirabilis
US-09-543-681A-5131

Query Match
Best Local Similarity 9.8%; Score 68.5; DB 4; Length 342;
Matches 25; Conservative 28; Mismatches 46; Indels 23; Gaps 5;

QY 2 TVPDN-----HRNKFVINVDGNEGLSGIMELT-----DTELILYTRKDSVK 46
DB 193 TIEDNPETFFHKVNYVHLFPSESDSKGSVVTPASLLRWYRNPDRILLTEVRGAE 252
QY 47 WHYLCRLRRYGYDSNLFSPESGRRCTQGQIFAF-----KCARAEELFNMQLQIMONNSIN 101
DB 253 WDFLKITGSGHEGSMSTIHAGSAKEAIDGFTRCYENPQCAQLPYTF-MLRKVL--DSDL 309
QY 102 VV 103
DB 310 VI 311

RESULT 14
US-08-894-017-23
; Sequence 23, Application US/08894017
; Patent No. 6024958
; GENERAL INFORMATION:
; APPLICANT: Lehner, Thomas
; APPLICANT: Kelly, Charles
; TITLE OF INVENTION: POLYPEPTIDE FRAGMENTS CAPABLE
; TITLE OF INVENTION: OF COMPETITION WITH STREPTOCOCCUS MUTANS ANTIGEN I/II
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESS: MORRISON & FOERSTER
; STREET: 2000 Pennsylvania Avenue, NW, suite 5500
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20006-1888
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows
; SOFTWARE: FASTSEQ for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/894.017
; FILING DATE: 31-JUL-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB96/00207
; FILING DATE: 31-JAN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Murashige, Kate H
; REGISTRATION NUMBER: 29,959
; REFERENCE/DOCKET NUMBER: 25150-20067.00
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-887-1500
; TELEFAX: 202-822-0168
; TELEX:
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1561 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-894-017-23

Query Match
Best Local Similarity 9.8%; Score 69.5; DB 3; Length 1561;
Matches 31; Conservative 18; Mismatches 58; Indels 45; Gaps 4;

QY 3 VPDNRNKFVINVDGNEGLSGIMELTDTELILY-----SGIMELTDTELILY----- 38
```

Db 1367 IPANHSEELFEYFYDDYDQDGHYTGQYKVFQKVDITLKNGVIIKSGTELTQYTTAEVD 1426
Qy 39 -TRKRDVSKWHYLCRLRYGDSNLFSESGRRCTGGIFAFKCARAEELFNMLQEIOMN 97
Db 1427 TTGATITKFEAFLSVSIDS-----AFQAESYIQMKRIAVGTFFN 1468
Qy 98 NSINNVVEEPVVERNNHOTELEVPRTPTTP 129
Db 1469 TYINTVNGVTYSSNTVKT--TTPEDPADPTDP 1498

RESULT 15
US-09-456-474-23
; Sequence 23, Application US/09456474
; Patent No. 6500433
; GENERAL INFORMATION:
; APPLICANT: Lehner, Thomas
; APPLICANT: Kelly, Charles
; TITLE OF INVENTION: POLYPEPTIDE FRAGMENTS CAPABLE OF
; FILE OF INVENTION: COMPETITION WITH STREPTOCOCCUS MUTANS ANTIGEN I/II
; FILE REFERENCE: 25150-20067.10
; CURRENT APPLICATION NUMBER: US/09/456,474
; CURRENT FILING DATE: 1999-12-09
; PRIOR APPLICATION NUMBER: US 08/894,017
; PRIOR FILING DATE: 1997-10-20
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 23
; LENGTH: 1561
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-09-456-474-23

Query Match 9.8%; Score 68.5; DB 4; Length 1561;
Best Local Similarity 20.4%; Pred. No. 34;
Matches 31; Conservative 18; Mismatches 58; Indels 45; Gaps 4;
Qy 3 VPDNERNKFVINVDGNELG-----SGIMELTDTLLIY----- 38
Db 1367 IPANHSEELFEYFYDDYDQDGHYTGQYKVFQKVDITLKNGVIIKSGTELTQYTTAEVD 1426
Qy 39 -TRKRDVSKWHYLCRLRYGDSNLFSESGRRCTGGIFAFKCARAEELFNMLQEIOMN 97
Db 1427 TTGATITKFEAFLSVSIDS-----AFQAESYIQMKRIAVGTFFN 1468
Qy 98 NSINNVVEEPVVERNNHOTELEVPRTPTTP 129
Db 1469 TYINTVNGVTYSSNTVKT--TTPEDPADPTDP 1498

Search completed: August 2, 2004, 09:32:28
Job time : 24.8026 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 2, 2004, 09:31:19 ; Search time 69.2763 Seconds
(without alignments)
588.640 Million cell updates/sec

Title: US-09-757-415A-1_COPY_11_140

Perfect score: 696

Sequence: 1 DTVPDNRHNRKFKVINDDG.....NNHQTELEVPRTPTTPG 130

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1291235 seqs, 313682936 residues

Total number of hits satisfying chosen parameters: 1291235

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:

```
1: /cgn2_6/prodata/2/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/prodata/2/pubpaa/PCT_NEW PUB.pep:*
3: /cgn2_6/prodata/2/pubpaa/US06_NEW PUB.pep:*
4: /cgn2_6/prodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/prodata/2/pubpaa/US07_NEW PUB.pep:*
6: /cgn2_6/prodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/prodata/2/pubpaa/US08_NEW PUB.pep:*
8: /cgn2_6/prodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/prodata/2/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/prodata/2/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/prodata/2/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/prodata/2/pubpaa/US09_NEW PUB.pep:*
13: /cgn2_6/prodata/2/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/prodata/2/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/prodata/2/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/prodata/2/pubpaa/US10_NEW PUB.pep:*
17: /cgn2_6/prodata/2/pubpaa/US60_NEW PUB.pep:*
18: /cgn2_6/prodata/2/pubpaa/US60_PUBCOMB.pep:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	696	100.0	508	10	US-09-757-415A-1
2	696	100.0	508	14	US-10-146-473-67
3	696	100.0	521	12	US-10-276-774-2192
4	685	98.4	508	12	US-09-731-660A-1
5	679	97.6	129	12	US-09-731-660A-3
6	143.5	20.6	541	13	US-10-001-870-179
7	141	20.3	359	9	US-09-789-519-58
8	137.5	19.8	331	15	US-10-307-928A-6
9	120.5	17.3	440	12	US-10-112-944-285
10	111.5	16.0	289	14	US-10-106-698-6260
11	95.5	13.7	268	14	US-10-029-386-33396
12	95.5	13.7	268	15	US-10-264-049-2340
13	78	11.2	105	14	US-10-192-381-5
14	78	11.2	114	12	US-09-731-660A-2
15	78	11.2	1231	16	US-10-694-874-3

16	78	11.2	1242	9	US-09-903-248-5	Sequence 5, Appli
17	78	11.2	1242	9	US-09-859-604-5	Sequence 5, Appli
18	78	11.2	1242	9	US-09-903-063-5	Sequence 5, Appli
19	78	11.2	1242	9	US-09-903-216-5	Sequence 5, Appli
20	78	11.2	1242	9	US-09-903-199-5	Sequence 5, Appli
21	78	11.2	1242	9	US-09-903-023-5	Sequence 5, Appli
22	78	11.2	1242	10	US-09-436-184-5	Sequence 5, Appli
23	78	11.2	1242	13	US-10-085-027-1	Sequence 1, Appli
24	78	11.2	1242	16	US-10-694-874-1	Sequence 1, Appli
25	78	11.2	1316	15	US-10-334-143-10	Sequence 10, Appl
26	75	10.8	99	12	US-10-424-599-217303	Sequence 217303,
27	75	10.8	537	13	US-10-047-542-74	Sequence 74, Appl
28	74.5	10.7	520	12	US-10-282-122A-71887	Sequence 71887, A
29	74.5	10.7	523	16	US-10-437-963-124637	Sequence 124637,
30	74.5	10.7	709	12	US-10-351-334-132	Sequence 132, App
31	74.5	10.7	973	12	US-10-276-774-2310	Sequence 2310, Ap
32	74	10.6	105	12	US-10-424-599-213391	Sequence 213391,
33	74	10.6	428	12	US-10-282-122A-45402	Sequence 45402, A
34	72.5	10.4	384	16	US-10-437-963-125379	Sequence 125379,
35	72.5	10.4	437	14	US-10-156-761-12680	Sequence 12680, A
36	72.5	10.4	501	14	US-10-029-386-32319	Sequence 32319, A
37	72.5	10.4	519	12	US-10-282-122A-71027	Sequence 71027, A
38	72	10.3	530	15	US-10-094-749-2723	Sequence 2723, Ap
39	72	10.3	1331	14	US-10-087-464-50	Sequence 50, Appl
40	71.5	10.3	1309	15	US-10-369-493-1656	Sequence 1656, Ap
41	71	10.2	516	12	US-10-282-122A-47351	Sequence 47351, A
42	71	10.2	633	12	US-10-425-114-43531	Sequence 43531, A
43	70.5	10.1	749	9	US-09-899-569A-2	Sequence 2, Appli
44	70.5	10.1	836	9	US-09-899-569A-4	Sequence 4, Appli
45	70.5	10.1	3170	14	US-10-128-714-8504	Sequence 8504, Ap

ALIGNMENTS

RESULT 1

US-09-757-415A-1
; Sequence 1, Application US/09757415A
; Publication No. US2003040612A1
; GENERAL INFORMATION:
; APPLICANT: Zhou, Ming-Ming
; TITLE OF INVENTION: Methods of Identifying Modulators of the FGF Receptor
; FILE REFERENCE: 2459-1-002N
; CURRENT APPLICATION NUMBER: US/09/757,415A
; CURRENT FILING DATE: 2001-03-09
; PRIOR APPLICATION NUMBER: 60/175967
; PRIOR FILING DATE: 2000-01-12
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 508
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-757-415A-1

Query Match	100.0%	Score 696;	DB 10;	Length 508;
Best Local Similarity	100.0%	Pred. No. 3.2e-73;		
Matches 130;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	DTVPDNRHNRKFKVINDDG	ELGSGIMELDT	TELLILYTRKEDSVKWHYLCRRYGYDSN 50
Db	11	DTVPDNRHNRKFKVINDDG	ELGSGIMELDT	TELLILYTRKEDSVKWHYLCRRYGYDSN 70
QY	61	LFSFSGRCOTGGGIFAFKCARAEELFNMLQEI	MONNSINVV	EEPPVVERNNHQTLEVP 120
Db	71	LFSFSGRCOTGGGIFAFKCARAEELFNMLQEI	MONNSINVV	EEPPVVERNNHQTLEVP 130
QY	121	RTPTPTTPG	130	
Db	131	RTPTPTTPG	140	

RESULT 2

```

US-10-146-473-67
; Sequence 67, Application US/10146473
; Publication No. US2003010888A1
; GENERAL INFORMATION:
; APPLICANT: Scanlan, Matthew
; APPLICANT: Gout, Ivan
; APPLICANT: Stockert, Elisabeth
; APPLICANT: Gure, Ali
; APPLICANT: Chen, Yao-Tseng
; APPLICANT: Old, Lloyd
; TITLE OF INVENTION: Breast Cancer Antigens
; FILE REFERENCE: L00461/70130(JRV)
; CURRENT APPLICATION NUMBER: US/10/146,473
; CURRENT FILING DATE: 2002-05-15
; PRIOR APPLICATION NUMBER: US 60/291,150
; PRIOR FILING DATE: 2001-05-15
; NUMBER OF SEQ ID NOS: 82
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 67
; LENGTH: 508
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-146-473-67

Query Match      100.0%; Score 696; DB 14; Length 508;
Best Local Similarity 100.0%; Pred. No. 3.2e-73;
Matches 130; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DTVPDHNKFKVINVDGNGELSGIMELTDTELLIYTRKDSVKWHYLCRLRYGYDSN 60
DB 11 DTVPDHNKFKVINVDGNGELSGIMELTDTELLIYTRKDSVKWHYLCRLRYGYDSN 70
QY 61 LFSFESGRCCTGGGIFAFKCARAEELFNMQLQEIOMNNSINVVEPVPVVRNNHOTELEVP 120
DB 71 LFSFESGRCCTGGGIFAFKCARAEELFNMQLQEIOMNNSINVVEPVPVVRNNHOTELEVP 130
QY 121 RTPRTPTTPG 130
DB 131 RTPRTPTTPG 140

RESULT 3
US-10-276-774-2192
; Sequence 2192, Application US/10276774
; Publication No. US20040053245A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; APPLICANT: Tang, Y. Tom et al
; TITLE OF INVENTION: No. US20040053245A1e1 Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-030
; CURRENT APPLICATION NUMBER: US/10/276,774
; CURRENT FILING DATE: 2002-11-18
; PRIOR APPLICATION NUMBER: 09/560,875
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 09/496,914
; PRIOR FILING DATE: 2000-02-03
; NUMBER OF SEQ ID NOS: 2700
; SOFTWARE: Custom
; SEQ ID NO 2192
; LENGTH: 521
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-276-774-2192

Query Match      100.0%; Score 696; DB 12; Length 521;
Best Local Similarity 100.0%; Pred. No. 3.4e-73;
Matches 130; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DTVPDHNKFKVINVDGNGELSGIMELTDTELLIYTRKDSVKWHYLCRLRYGYDSN 60
DB 24 DTVPDHNKFKVINVDGNGELSGIMELTDTELLIYTRKDSVKWHYLCRLRYGYDSN 83
QY 61 LFSFESGRCCTGGGIFAFKCARAEELFNMQLQEIOMNNSINVVEPVPVVRNNHOTELEVP 120

```

```

US-09-731-660A-1
; Sequence 1, Application US/09731660A
; Publication No. US20020086972A1
; GENERAL INFORMATION:
; APPLICANT: KOUHARA, HARUHIKO
; APPLICANT: SPIVAK-KROIZMAN, TALY
; APPLICANT: LAX, IRIT
; APPLICANT: SCHLESSINGER, JOSEPH
; TITLE OF INVENTION: ADAPTOR PROTEIN FR2 AND RELATED PRODUCTS AND METHODS
; FILE REFERENCE: 038602/1023
; CURRENT APPLICATION NUMBER: US/09/731,660A
; CURRENT FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 08/980,523
; PRIOR FILING DATE: 1997-12-01
; PRIOR APPLICATION NUMBER: 60/032,093
; PRIOR FILING DATE: 1996-12-03
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 1
; LENGTH: 508
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-731-660A-1

Query Match      98.4%; Score 685; DB 12; Length 508;
Best Local Similarity 97.7%; Pred. No. 6.5e-72;
Matches 127; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 DTVPDHNKFKVINVDGNGELSGIMELTDTELLIYTRKDSVKWHYLCRLRYGYDSN 60
DB 11 DTVPDHNKFKVINVDGNGELSGIMELTDTELLIYTRKDSVKWHYLCRLRYGYDSN 70
QY 61 LFSFESGRCCTGGGIFAFKCARAEELFNMQLQEIOMNNSINVVEPVPVVRNNHOTELEVP 120
DB 71 LFSFESGRCCTGGGIFAFKCARAEELFNMQLQEIOMNNSINVVEPVPVVRNNHOTELEVP 130
QY 121 RTPRTPTTPG 130
DB 131 RTPRTPTTPG 140

RESULT 5
US-09-731-660A-3
; Sequence 3, Application US/09731660A
; Publication No. US20020086972A1
; GENERAL INFORMATION:
; APPLICANT: KOUHARA, HARUHIKO
; APPLICANT: SPIVAK-KROIZMAN, TALY
; APPLICANT: LAX, IRIT
; APPLICANT: SCHLESSINGER, JOSEPH
; TITLE OF INVENTION: ADAPTOR PROTEIN FR2 AND RELATED PRODUCTS AND METHODS
; FILE REFERENCE: 038602/1023
; CURRENT APPLICATION NUMBER: US/09/731,660A
; CURRENT FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 08/980,523
; PRIOR FILING DATE: 1997-12-01
; PRIOR APPLICATION NUMBER: 60/032,093
; PRIOR FILING DATE: 1996-12-03
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 3
; LENGTH: 129
; TYPE: PRT
; ORGANISM: Homo sapiens

```

```
US-09-731-660A-3
Query Match          97.6%; Score 679; DB 12; Length 129;
Best Local Similarity 97.7%; Pred. No. 5.1e-72;
Matches 126; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 DTVPDNRHNRKFKVINYDDNGELGSGIMELTDTLLYTRKDSVKWHYLCRRYGYDSN 60
DB 1 DTVPDNRHNRKFKVINYDDNGELGSGIMELTDTLLYTRKDSVKWHYLCRRYGYDSN 60
QY 61 LFSFESGRRRCQTGGIFAFKCARAEELFNNLMQIMQNSINVVEPVRNHNHOTELEVP 120
DB 61 LFSFESGRRRCQTGGIFAFKCARAEELFNNLMQIMQNSINVVEPVRNHNHOTELEVP 120
QY 121 RTPTPTTP 129
DB 121 RTPTPTTP 129

RESULT 6
US-10-001-870-179
; Sequence 179, Application US/10001870
; Publication No. US20020150924A1
; GENERAL INFORMATION:
; APPLICANT: Salceda, Susana
; APPLICANT: Macina, Roberto
; APPLICANT: Recipon, Hervé
; APPLICANT: Sun, Yongming
; APPLICANT: Liu, Chenghua
; TITLE OF INVENTION: Compositions and Methods Relating to Prostate Specific Genes and
; FILE REFERENCE: DEX-0283
; CURRENT APPLICATION NUMBER: US/10/001,870
; CURRENT FILING DATE: 2001-11-20
; PRIOR APPLICATION NUMBER: 60/252,189
; PRIOR FILING DATE: 2000-11-21
; NUMBER OF SEQ ID NOS: 217
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 179
; LENGTH: 541
; TYPE: PRT
; ORGANISM: Homo sapien
US-10-001-870-179

Query Match          20.6%; Score 143.5; DB 13; Length 541;
Best Local Similarity 29.5%; Pred. No. 7.4e-08;
Matches 33; Conservative 20; Mismatches 50; Indels 9; Gaps 2;

QY 24 GSGIMELTDTLLYTRKDS-----VKWHYLCRRYGYDSNLFSPESGRRRCQTGGIF 77
DB 248 GSYVLRVEAERLTLLTVGAGSQILEPLLSWPTLLRYGRDKYMFSEAGRCPSGPGTF 307
QY 78 AFKCARAEELFNNLMQIMQNSINVVE---EPVVRNHNHOTELEVPRTPTTP 126
DB 308 TFTAQNDIFQAVETAIRHQKAGQAGQGHVLRADSHGEVAEGKLPSP 359

RESULT 7
US-09-789-919-59
; Sequence 58, Application US/09789919
; Patent No. US20020064855A1
; GENERAL INFORMATION:
; APPLICANT: Lemischka, Ihor
; APPLICANT: Moore, Kateri
; TITLE OF INVENTION: GENES THAT REGULATE HEMATOPOIETIC BLOOD FORMING STEM
; FILE REFERENCE: 2275-1-005
; CURRENT APPLICATION NUMBER: US/09/789,919
; CURRENT FILING DATE: 2001-02-21
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 58
; LENGTH: 359
; TYPE: PRT
```

```
; ORGANISM: Mus musculus
US-09-789-919-58

Query Match          20.3%; Score 141; DB 9; Length 359;
Best Local Similarity 28.3%; Pred. No. 8.4e-08;
Matches 36; Conservative 20; Mismatches 47; Indels 24; Gaps 3;

QY 23 LGSIMELTDTLLYTRKDSVKWHYLCRRYGYDSNLFSPESGRRRCQTGGIFAFKCA 82
DB 184 LGQDDIQURET-----SKPQACFPWPFRFLKYGSKGVFSFAGRRCDSEGLFAFSSP 238
QY 83 RAEELFNNLMQIMQNSINVVE-----EPVVRNHNHOTELEVPRTPTTP 124
DB 239 RAPDICGVAAAIARQRELPELAMSPPCPLRALSLPSLEPPGPELREAVAPGFLPTPRK 298
QY 125 TP-TTPG 130
DB 299 LPLTDPG 305

RESULT 8
US-10-307-928A-6
; Sequence 6, Application US/10307928A
; Publication No. US20030229016A1
; GENERAL INFORMATION:
; APPLICANT: Alsobrook, John P.
; APPLICANT: Anderson, David W.
; APPLICANT: Boldog, Ferenc L.
; APPLICANT: Burgess, Catherine E.
; APPLICANT: Catterton, Elina
; APPLICANT: Edinger, Shlomit R.
; APPLICANT: Gorman, Linda
; APPLICANT: Guo, Xiaojia (Sasha)
; APPLICANT: Ji, Weizhen
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Li, Li
; APPLICANT: Patturajan, Meera
; APPLICANT: Rieger, Daniel K.
; APPLICANT: Shenoy, Suresh G.
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Vernet, Corine A.M.
; APPLICANT: Voss, Edward Z.
; APPLICANT: Zhong, Mei
; TITLE OF INVENTION: NOVEL HUMAN PROTEINS, POLYNUCLEOTIDES ENCODING THEM AND METHODS OF
; FILE REFERENCE: 24102-502D
; CURRENT APPLICATION NUMBER: US/10/307,928A
; CURRENT FILING DATE: 2002-12-02
; PRIOR APPLICATION NUMBER: 60/341,477
; PRIOR FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: 60/341,540
; PRIOR FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: 60/342,592
; PRIOR FILING DATE: 2001-12-20
; PRIOR APPLICATION NUMBER: 60/344,903
; PRIOR FILING DATE: 2001-12-31
; PRIOR APPLICATION NUMBER: 60/373,288
; PRIOR FILING DATE: 2002-04-17
; PRIOR APPLICATION NUMBER: 60/380,981
; PRIOR FILING DATE: 2002-05-15
; PRIOR APPLICATION NUMBER: 60/381,495
; PRIOR FILING DATE: 2002-05-17
; PRIOR APPLICATION NUMBER: 60/383,744
; PRIOR FILING DATE: 2002-05-28
; PRIOR APPLICATION NUMBER: 60/384,024
; PRIOR FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: 60/401,788
; PRIOR FILING DATE: 2002-08-07
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: Curaseqlist version 0.1
; SEQ ID NO 6
; LENGTH: 331
```

```
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-307-928A-6

Query Match      19.8%; Score 137.5; DB 15; Length 331;
Best Local Similarity 31.5%; Pred. No. 1.9e-07;
Matches 34; Conservative 19; Mismatches 42; Indels 13; Gaps 3;

QY 24 GSGIMELTDELILYTRKRDVSK---WHYLCRLRYGYDSNLFSESGRRRCQTGGIFAFK 80
Db 153 GECTWQITHENILYLDIHNKVLWPLSSRLRYGRDSTWTFESGRMCDTGSLTTFQ 212
QY 81 CARAEELFNMLQEIOMNNSINVVEE-----PVVERNNHTELEVPRT 122
Db 213 TREGE---MYQKVHSATLAEQHERLMLEMQKARLQTSLEPMT 256

RESULT 9
US-10-112-944-285
; Sequence 285, Application US/10112944
; Publication No. US20040048249A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Yang, Yongchong
; APPLICANT: Weng, Gezhi
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Xue, Aidong J.
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Wehrman, Tom
; APPLICANT: Ghosh, Malabika
; APPLICANT: Wang, Dunrui
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wang, Zhiwei
; TITLE OF INVENTION: No. US20040048249A1el Nucleic Acids and
; FILE REFERENCE: 805A
; CURRENT APPLICATION NUMBER: US/10/112,944
; CURRENT FILING DATE: 2002-03-28
; PRIOR APPLICATION NUMBER: US 09/488,725
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US 09/491,404
; PRIOR FILING DATE: 2000-01-25
; PRIOR APPLICATION NUMBER: US 09/496,914
; PRIOR FILING DATE: 2000-02-03
; PRIOR APPLICATION NUMBER: US 09/515,126
; PRIOR FILING DATE: 2000-02-28
; PRIOR APPLICATION NUMBER: US 09/519,705
; PRIOR FILING DATE: 2000-03-07
; PRIOR APPLICATION NUMBER: US 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: US 09/552,929
; PRIOR FILING DATE: 2000-04-18
; PRIOR APPLICATION NUMBER: US 09/577,408
; PRIOR FILING DATE: 2000-05-18
; NUMBER OF SEQ ID NOS: 924
; SOFTWARE: pt_FL_genes Version 5.0
; SEQ ID NO 285
; LENGTH: 440
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-112-944-285

Query Match      17.3%; Score 120.5; DB 12; Length 440;
Best Local Similarity 29.7%; Pred. No. 2.9e-05;
Matches 35; Conservative 16; Mismatches 38; Indels 29; Gaps 5;

QY 23 LGSGIMELTTE--LILYTRKRSVKWHYLCRLRYGYDSNLFSESGRRRCQTGGIFAFK 80
Db 184 LGPDAIQLRKAGTQALYS-----WPYHFLRLKLGSDKGVFSFAGRRCHSGEGLF 236
QY 81 CARAEELFNMLQEIOMNNSINVVEEPVVERNNHTELE-----EVPTRPTPT--TPG 130

; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-106-698-6260
; Sequence 6260, Application US/10106698
; Publication No. US20030109690A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptides
; FILE REFERENCE: PA005PI
; CURRENT APPLICATION NUMBER: US/10/106,698
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: PCT/US00/26524
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US 60/157,137
; PRIOR FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: US 60/163,280
; PRIOR FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 8564
; SOFTWARE: PatentIn Ver. 3.0
; SEQ ID NO 6260
; LENGTH: 289
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (225)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: MISC FEATURE
; LOCATION: (228)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: MISC FEATURE
; LOCATION: (229)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-106-698-6260

Query Match      16.0%; Score 111.5; DB 14; Length 289;
Best Local Similarity 31.2%; Pred. No. 0.00019;
Matches 20; Conservative 15; Mismatches 26; Indels 3; Gaps 1;

QY 28 MELTDELILYTRKRDVSK---WHYLCRLRYGYDSNLFSESGRRRCQTGGIFAFKCARA 84
Db 202 LQITHENILYLDIHNPRVKLVSWKLCXXRYGRDTRTFEAGRMCDAGECLYTFQTQEG 261
QY 85 BELF 88
Db 262 EQIY 265

RESULT 11
US-10-029-386-33396
; Sequence 33396, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR GI
; FILE REFERENCE: AEONICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Annonax Sequence Listing Engine vers. 1.1
; SEQ ID NO 33396
; LENGTH: 268
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC006544.19
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.77
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1
```

```

; OTHER INFORMATION: SWISSPROT HIT: P56945, EVALU6 2.00e-03
US-10-029-386-33396

Query Match      13.7%; Score 95.5; DB 14; Length 268;
Best Local Similarity 27.5%; Pred. No. 0.013;
Matches 19; Conservative 16; Mismatches 31; Indels 3; Gaps 1;

Qy 61 LFSFESGRRCQTGGCIGAFKCARAEELFMLOETMONNSINVVE---EPVVERNNHOTEL 117
Db 2 MFSFEAGRCRCPSGPGTFTFQAQGNIDIFQAVETAHROKQAGQGHDVLRADSHGEV 61

Qy 118 EVPRTERTP 126
Db 62 AEGKLPSPP 70

RESULT 12
US-10-264-049-2340
; Sequence 2340, Application US/10264049
; Publication No. US20040005579A1
; GENERAL INFORMATION:
; APPLICANT: Birse et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA133P1
; CURRENT APPLICATION NUMBER: US/10/264,049
; CURRENT FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: PCT/US01/18569
; PRIOR FILING DATE: 2001-06-07
; PRIOR APPLICATION NUMBER: US 60/209,467
; PRIOR FILING DATE: 2000-06-07
; NUMBER OF SEQ ID NOS: 4360
; SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 2340
; LENGTH: 268
; TYP2: PRT
; ORGANISM: Homo sapiens
US-10-264-049-2340

Query Match      13.7%; Score 95.5; DB 15; Length 268;
Best Local Similarity 27.5%; Pred. No. 0.013;
Matches 19; Conservative 16; Mismatches 31; Indels 3; Gaps 1;

Qy 61 LFSFESGRRCQTGGCIGAFKCARAEELFMLOETMONNSINVVE---EPVVERNNHOTEL 117
Db 2 MFSFEAGRCRCPSGPGTFTFQAQGNIDIFQAVETAHROKQAGQGHDVLRADSHGEV 61

Qy 118 EVPRTERTP 126
Db 62 AEGKLPSPP 70

RESULT 13
US-10-192-381-5
; Sequence 5, Application US/10192381
; Publication No. US20030170807A1
; GENERAL INFORMATION:
; APPLICANT: THE JOHNS HOPKINS UNIVERSITY SCHOOL OF MEDICINE
; APPLICANT: WORLEY, Paul
; APPLICANT: TU, Jian
; APPLICANT: XIAO, Bo
; APPLICANT: LEAHY, Daniel
; APPLICANT: BENKEN, Jutta
; APPLICANT: LANAHAN, Anthony
; TITLE OF INVENTION: NUCLEIC ACID MOLECULE ENCODING HOMER 1b PROTEIN (AS
; FILE OF INVENTION: AMENDED)
; FILE REFERENCE: JHU1580-4
; CURRENT APPLICATION NUMBER: US/10/192,381
; CURRENT FILING DATE: 2002-07-09
; PRIOR APPLICATION NUMBER: US/09/377,285
; PRIOR FILING DATE: 1999-08-18
; PRIOR APPLICATION NUMBER: US 60/138,426
; PRIOR FILING DATE: 1999-06-10
; PRIOR APPLICATION NUMBER: US 60/138,493

```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 8, 2004, 19:45:10 ; Search time 4766.38 Seconds
(without alignments)
1182.153 Million cell updates/sec

Title: US-09-757-415a-1_COPY_11_140

Perfect score: 696

Sequence: 1 DTVPDNRKPKVINVDDG.....NNHQTELEVPRTPRTPTTPG 130

Scoring table:
BLOSUM62 Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DBV=xlh
-Q=/cgm2/USPTO.spool/US09757415/runat_02082004_101230_6685/app_query.fasta_1.526
-DB=GenEmbl -Qfmt=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORW=ext -HEADSIZE=570 -MINLEN=0 -MAXLEN=200000000
-USER=US09757415 @CGN 1 1 3851 @runat_02082004_101230_6685 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.*

1: gb_ba.*
2: gb_atg.*
3: gb_in.*
4: gb_ov.*
5: gb_ov.*
6: gb_ov.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vi.*
15: em_ba.*
16: em_funi.*
17: em_hum.*
18: em_in.*
19: em_mu.*
20: em_or.*
21: em_or.*
22: em_ov.*
23: em_pat.*
24: em_ph.*
25: em_pl.*
26: em_ro.*
27: em_sts.*
28: em_un.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	696	100.0	1532	9	AF036717 Homo sapi
2	696	100.0	1840	9	BC021562 Homo sapi
3	685	98.4	5701	10	BC043109 Mus muscu
4	685	98.4	5792	10	BC055334 Mus muscu
5	616.5	88.6	2280	5	AB064525 Xenopus l
6	616.5	88.6	3017	5	BC046943 Xenopus l
7	613.5	88.1	1612	5	AF390895 Xenopus l
8	517.5	74.4	766	6	BD079340 Cancer-as
9	510.5	73.3	2031	9	BC010611 Homo sapi
10	510.5	73.3	2077	9	AF036718 Homo sapi
11	484.5	69.6	2150	10	BC014819 Mus muscu
12	452	64.9	128252	9	U91327 Human chrom
13	452	64.9	188666	9	AC018921 Homo sapi
14	409.5	58.8	241432	10	AL589661 Mouse DNA
15	339.5	48.8	238379	2	AC106185 Rattus no
16	339.5	48.8	239624	2	AC120279 Rattus no
17	339.5	48.8	247654	2	AC122658 Rattus no
18	318.5	45.8	2531	5	BC060017 Xenopus l
19	297.5	42.7	136514	2	BC060017 Xenopus l
20	278	39.9	79288	2	BC060017 Xenopus l
21	272	39.1	106711	9	BC060017 Xenopus l
22	271.5	39.0	6834	10	AF456480 Human DNA
23	271.5	39.0	175795	2	AC124830 Mus muscu
24	268	38.5	189649	2	AC139641 Rattus no
25	268	38.5	290773	2	AC129162 Rattus no
26	182	26.1	189649	2	AC139641 Rattus no
27	182	26.1	298853	2	AC122659 Rattus no
28	174	25.0	2779	3	AY118363 Drosophil
29	174	25.0	36890	2	AC018157 Drosophil
30	174	25.0	79076	3	AC004423 Drosophil
31	174	25.0	186490	3	AC092234 Drosophil
32	174	25.0	308031	3	AE003821 Drosophil
33	151.5	21.8	38094	3	AF040847 Caenorhab
34	151.5	21.8	278007	2	AC006799 Caenorhab
35	145	20.8	1749	10	MMU78818 Mus muscu
36	145	20.8	1812	10	BC013066 Mus muscu
37	143.5	20.6	1873	9	HSU70987 Human GAP b
38	143.5	20.6	1972	9	AF035299 Homo sapi
39	143.5	20.6	2473	9	AK055944 Homo sapi
40	143.5	20.6	3809	6	AX478178 Sequence
41	142	20.4	205	6	BD071770 Secreted
42	142	20.4	205	6	BD072054 Secreted
43	141	20.3	1535	10	AF237580 Mus muscu
44	141	20.3	1539	6	AX046348 Sequence
45	141	20.3	1565	10	AF179242 Mus muscu

ALIGNMENTS

RESULT 1

AF036717	1532 bp	mRNA	linear	PRI 30-JUL-1999
LOCUS				
DEFINITION	Homo sapiens FGFR signalling adaptor SNT-1 mRNA, complete cds.			
ACCESSION	AF036717			
VERSION	J. Biol. Chem. 273 (29), 17987-17990 (1998)			
KEYWORDS	AF036717.1 GI:2708627			
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens			
REFERENCE	1 (bases 1 to 1532)			
AUTHORS	Xu, H., Lee, K.W. and Goldfarb, M.			
TITLE	Novel recognition motif on fibroblast growth factor receptor mediates direct association and activation of SNT adaptor proteins			
JOURNAL	J. Biol. Chem. 273 (29), 17987-17990 (1998)			
MEDLINE	98324992			
PUBMED	9660748			
REFERENCE	2 (bases 1 to 1532)			
AUTHORS	Xu, H., Lee, K. and Goldfarb, M.			
TITLE	Direct Submission			
JOURNAL	Submitted (04-DEC-1997) Brookdale Center for Developmental and Molecular Biology, Mt. Sinai School of Medicine, 1 Gustave Levy Place, New York, NY 10029, USA			
FEATURES	Location/Qualifiers			
source	1..1532			
	/organism="Homo sapiens"			
	/mol_type="mRNA"			
	/db_xref="taxon:9606"			
	/chromosome="12"			
	/map="12p15"			
	/tissue type="placenta"			
gene	1..1532			
	/gene="SNT-1"			
CDS	6..1532			
	/note="similar to murine FR2"			
	/codon_start=1			
	/product="FGFR signalling adaptor SNT-1"			
	/protein_id="AAB92554.1"			
	/db_xref="GI:2708628"			
	/translation="MGSCCSPDKDTPVDRHNRKFKVNVDDCNELGGIMELTDTLILYTRKSDVKHYLCRRGYDNLFSFSGRRCTGGIGFAFKARAEELFNMLQEIQNNINVEEPVNRNHHQTELEVPRTPTTGFPAALPNQYRYPSPFGDASHPSSRPSVGSARLPVSGSESTHPLVASEQVHTVYNTTGVQEEKNRTSVHVPLEA RVSNAGSTPKPEESIEDRPQILLPEGVKFLVPTPVQKLMKEKLEQLGRQVSGCANNTENDTGYSDSDERDAPSVNKLVENINGLSIPASGVRRGLRTSTSTDTQ NINNSAQRRTALAYENLPSLPVWEARKLSRDEDNLGKPTSLANGYHNNLDPMHNY VNTENVTPASAHKIEYSRRDCPTVFNFDIRPSLEHQLNLYIQVDLEGSDSDNP QTPKTPITPQPTTRTELYAVDIERTAMSNLQALPRDDGTGRKTRHNSDLPDM"			
	6..23			
misc_feature	/gene="SNT-1"			
	/note="encodes myristoylation motif"			
misc_feature	51..398			
	/gene="SNT-1"			
	/note="encodes PTB domain"			
ORIGIN				
Alignment Scores:				
Pred. No.:	1..92e-79	Length:	1532	
Score:	696.00	Matches:	130	
Percent Similarity:	100.00%	Conservative:	0	
Best Local Similarity:	100.00%	Mismatches:	0	
Query Match:	100.00%	Indels:	0	
DB:	9	Gaps:	0	
US-09-757-415A-1_COPY_11_140 (1-130) x AF036717 (1-1532)				
Qy	1	AspThrValProAspAsnHisArgAsnLysPheLysValIleAsnValAspAspGly	20	
Db	36	GACACTGCCAGATAACCATCGGAACAGTTAAGTCATTAATGTTGGATGATGGG	95	
Qy	21	AsnGluLeuGlySerGlyIleMetGluLeuThrAspThrGluLeuLeuLeuThrArg	40	
Db	96	AATGAGTTAGGTTCTGGCATATGGAACCTTACAGACACAGAACTGATTTTATACACCGC	155	
Qy	41	LysArgAspSerValLysTrpHisTyrLeuCysLeuArgArgTyrGlyTyrAspSerAsn	60	
Db	156	AAACGTGACTCACTAATAATGGCACTACCTTGCCTGCGACGCTATGGTATGATCTCGAT	215	
Qy	61	LeuPheSerPheGluSerGlyArgArgCysGlnThrGlyGlnGlyIlePheAlaPheLys	80	
Db	216	CTCTTTTCTTTTGAAGTGGTCAAGAGTGTCAAACTGCACAGGAATCTTTCCTTTAAG	275	
Qy	81	CysAlaArgAlaGluLeuPheAsnMetLeuGlnGluIleMetGlnAsnAsnSerIle	100	
Db	276	TGTGCCGTGCGAAGAATTTATTAACTGTTGCAAGAGATTATGCAAAATAATAGTATA	335	
Qy	101	AsnValValGluGluProValValGluArgAsnAsnHisGlnThrGluLeuGluValPro	120	
Db	336	AATGTTGTGGAAGAGCCAGTTGTGAGAAAGAAATAATCATCAGACAGAAATTCGAAGTCCT	395	
Qy	121	ArgThrProArgThrProThrThrProGly	130	
Db	396	AGAACACCTCGAACACCTCAACTCCAGGA	425	
RESULT 2				
LOCUS	BC021562 1840 bp mRNA linear PRI 06-OCT-2003			
DEFINITION	Homo sapiens fibroblast growth factor receptor substrate 2, mRNA (CDNA clone MGC:31881 IMAGE:4556225), complete cds.			
ACCESSION	BC021562			
VERSION	BC021562.1 GI:18204300			
KEYWORDS	MGC.			
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens			
REFERENCE	1 (bases 1 to 1840)			
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
	Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.P., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carrinci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullaly, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalob, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smal, U., Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.			
	Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences			
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)			
MEDLINE	22388257			
PUBMED	12477932			
REFERENCE	2 (bases 1 to 1840)			
AUTHORS	Straussberg, R.			
TITLE	Direct Submission			
JOURNAL	Submitted (14-JAN-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA			
REMARK	NIH-MGC Project URL: http://mgc.nci.nih.gov			
COMMENT	Contact: MGC help desk Email: cgapbs-remail.nih.gov Tissue Procurement: ATCC CDNA Library Preparation: Rubin Laboratory CDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL) DNA Sequencing by: Genome Sequence Centre,			

BC Cancer Agency, Vancouver, BC, Canada
 info@bcsc.bc.ca
 Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,
 Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,
 Letitia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo
 Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven
 Ness, Pawan Pandoh, Anna-Liisa Prabhu, Parvaneh Saeedi, Jacqueline
 Schein, Duane Smalls, Michael Smith, Lorraine Spence, Jeff Stott,
 Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy,
 George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAL Plate: 40 Row: e Column: 6
 This clone was selected for full length sequencing because it
 passed the following selection criteria: matched mRNA gi: 21314643.
 Location/Qualifiers
 1. .1840
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="MGC:31831 IMAGE:4556225"
 /tissue_type="Uterus, leiomyosarcoma"
 /clone_lib="NIH_MGC_46"
 /lab_host="DH10B-R"
 /notes="Vector: pOTB7"
 1. .1840
 /gene="FRS2"
 /note="synonyms: SNT-1, FRS2alpha, SNT1, FRS2A, SNT"
 /db_xref="LocusID:10818"
 /db_xref="MIM:607743"
 247. .1785
 /codon_start=1
 /product="sucl-1 associated neurotrophic factor target (FGFR
 signalling adapt"
 /protein_id="AA421562.1"
 /db_xref="GI:18204301"
 /db_xref="LocusID:10818"
 /translations="MGSCSPDKDTVPDNRHNRKFKVINVDGNEGLSGIMELDTDE
 LILYTKRSDKWHYICLARYGDSNLFESGRCTQGGIFAFKCAAEFLNMLQ
 EIMONNSINVEPVRVNRNHOETELVPRTPTPTGPAQNLNPNRYSPFSGDAS
 SHPSRPSVSARLPVGEESTHLLVAEEQVHTVNTTGVQEKRNKTSVHVPLEA
 RVSNASTTPKPSIEDRDPOILLEPESGKVLGPTPVOKLMEKLEOLGRDOV
 SGGANNTWDQTDSDRDAEVSUNKLYENINGLSIPASGVRGRLTSTSDTQ
 NINSAQRVALLNLSPLPPWEARKLSDDEDNLGKPTSLNGYHNLDPHNY
 VNTENTVPASAKIKYSRRDCTPTVFNDIRPSLEHRLNLIQVDLEGSDSDNP
 QTEKPTPTLPQTPTARTELYAVIDIERTAAASNLOKALPRDDGTGKTRHNSIDLPM
 LAWK"
 298. 576
 /note="IRS; Region: PTB domain (IRS-1 type)";
 /db_xref="CDD:pfam02174"

misc_feature
 298. 576
 /note="IRS; Region: PTB domain (IRS-1 type)";
 /db_xref="CDD:pfam02174"

ORIGIN

Alignment Scores:
 Pred. No.: 2,45e-79 Length: 1840
 Score: 696.00 Matches: 130
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 9 Gaps: 0

US-09-757-415A-1_COPY_11_140 (1-130) x BC021562 (1-1840)

Qy 1 AspThrValProAspAsnHisArgAsnLysPheLysValIleAsnValAspAspGly 20
 |||||
 Db 277 GACACTGCCAGATACCAACGCAAGTTTAAGTCAATTATGTCGATGATGCGG 336
 |||||

Qy 21 AsnGluLeuGlySerGlyIleMetGluLeuThrAspThrGluLeuIleLeuTyrThrArg 40
 |||||
 Db 337 AATGAGTTAGTTCTGCAATAATGAACTTACACACAGACTGATTTATACCCGC 396
 |||||

Qy 41 LysArgAspSerValLysTyrHisTyrLeuCysLeuArgArgTyrGlyTyrAspSerAsn 60
 |||||

Db 397 AACGTGACTCAGTAAATGGCACTACTCTGCTGCGAAGCTATGCTGACTCGAAT 456
 |||||

Qy 51 LeuPheSerPheGluSerGlyArgCysGlnThrGlyGlnGlyIlePheAlaPheLys 80
 |||||

Db 457 CTCCTTTCTTTTGAAGTGGTGGGAGGTGCTCAACTGACAGAACTTTCCTTTTAA 516
 |||||

Qy 81 CysAlaArgAlaGluGluLeuPheAsnMetLeuGlnGluIleMetGlnAsnSerIle 100
 |||||

Db 517 TGTGCCGTGCAGAGAATATTATTAAACATGTTGCAAGAGATTATGCAAAATAATAGTATA 576
 |||||

Qy 101 AsnValValGluGluProValValGluArgAsnHisGlnThrGluLeuGluValPro 120
 |||||

Db 577 AATGTGGTGGAGAGCCAGTGTGTAGAAGAAATAATCATCAGACAGAAATGGAAGTCCCT 636
 |||||

Qy 121 ArgThrProArgThrProThrThrProGly 130
 |||||

Db 637 AGAACACCTCGACACCTTACACTCCAGGA 666
 |||||

RESULT 3
 BC043109 5701 bp mRNA linear ROD 11-DEC-2003
 LOCUS Mus musculus fibroblast growth factor receptor substrate 2, mRNA
 DEFINITION (cDNA clone MGC:58030 IMAGE:6406912), complete cds.
 ACCESSION BC043109.1 GI:27695419
 VERSION BC043109.1
 KEYWORDS MGC.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 5701)
 Strausberg,R.H., Feingold,B.A., Grouse,L.H., Derge,J.G.,
 Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
 Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
 Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
 Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
 Diatleron,K., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
 Schett,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
 Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
 Abramson,R.D., Mullany,S.J., Bosak,S.A., McEwan,P.J.,
 McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
 Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
 Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
 Fahey,J., Helton,E., Kettman,M., Madan,A., Young,A.C., Shetchenko,Y.,
 Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shetchenko,Y.,
 Bouffard,G.G., Blakesley,R.W., Grimwood,J.W., Green,E.D.,
 Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
 Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smalls,D.E.,
 Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
 Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
 12477932
 2 (bases 1 to 5701)
 Strausberg,R.
 Direct Submission
 Submitted (09-JAN-2003) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgabbs-r@mail.nih.gov
 Tissue Procurement: Dr. Jim Lin, University of Iowa
 cDNA Library Preparation: M. Bento Soares, University of Iowa
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: University of Iowa, Dr. M. Bento Soares and Dr.
 Thomas L. Casavant
 Web site: <http://genome.uiowa.edu>
 Contact: bento-soares@uiowa.edu; tom-casavant@uiowa.edu
 Ronaldo,M.F., Akabogu,I., Bair,T., Bair,J., Crouch,K., Davis,A.,
 Fishler,K., Keppel,C., Kucaba,T., Lebeck,M., Melo,A., Schaefer,K.,
 Schletz,I., Smith,C., Snir,E., Tack,D., Trout,K., Walters,J.,

Casavant, T., Soares, M.B.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: Plate: Row: Column: 0
This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis, Genomescan gene prediction.

FEATURES

Location/Qualifiers
1..5701
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="MGC:58030 IMAGE:6406912"
/tissue_type="Brain, enriched mouse brain 12.5dp"
/clone_lib="NIH_BMAP_F00"
/lab_host="DH10B"
/note="vector: pYX-ASC"

gene

1..5701
/gene="Frs2"
/note="synonyms: 4732458E18, SNT1"
/db_xref="LOCUSID:327826"
356..1882
/codon_start=1
/product="Frs2 protein"
/protein_id="AAH43109.1"
/db_xref="GI:27695420"
/db_xref="LOCUSID:327826"

CDS

/translation="MGSCCSPKDPDTPDNRNFKVINVDDNGLGSGWELTDT
LILTRKDSVKHYLCVDSNLFESGRCQTGGIYFAFKARAEFLNMLQ
EIMQNSINVEEVPWERSHTELEVPRTPTTGLCAQLPNRYPSFGDAS
SHPSRNPVSGAALPVGSESTPLVAEEQVHTVNTTGVQEEKRNASVHPPEA
RVSAENETKPEBPEDDPVLLKPGVRFVGLPTPVQKLMKEKLEQLGKDPV
SGSCAGNTWDGVDSDERDVPVKNLVENGLSIPSAGVRGRGLTSTSDTQ
NINNSARRALLNENLPSLPVWEARKLSRDEDDNLGKPSLNGYHNLDPMNY
VNTENVTPSAHKIDYSKRDCTVFNFDIRPSLEHRLNLYIQVDLEGGSDSNP
QXTEPTPLPQPTRETELYAVIDIBRTAAMNLOKALPRDDGTGRKTRHNSDLP
M"
407..685
/note="IRS, Region: PTB domain (IRS-1 type)"
/db_xref="CDD:pfam02174"

misc_feature

407..685
/note="IRS, Region: PTB domain (IRS-1 type)"
/db_xref="CDD:pfam02174"

ORIGIN

Alignment Scores:

Pred. No.: 2,86e-77 Length: 5701
Score: 685.00 Matches: 127
Percent Similarity: 100.00% Conservative: 3
Best Local Similarity: 97.59% Mismatches: 0
Query Match: 98.42% Indels: 0
DB: 10 Gaps: 0

US-09-757-415A-1_COPY_11_140 (1-130) x BC043109 (1-5701)

Qy	1	AspThrValProAspAsnHisArgAsnLysPheLysValIleAsnValAspAspAspGly	20
Db	386	GACACTGCCAGATACCATCGAACAGATTAAAGTCATTATGTGGATGATGGG	445
Qy	21	AsnGluLeuGlySerGlyIleMetGluLeuThrAspThrGluLeuLeuLeuThrArg	40
Db	446	AATGAGCTAGGCTCTGGTGTGATGGAACACACACAGAGCTGATCTGTACACCCGG	505
Qy	41	LysArgAspSerValLysTrpHisTyrLeuCysLeuArgArgTyrGlyTyrAspSerAsn	60
Db	506	AAACGGGACTCGGTAATGCATCCTCTGCTACGACATACGCTATGACTCAAT	565
Qy	61	LeuPheSerPheGluSerGlyArgArgCysGlnThrGlyGlnGlyIlePheAlaPheLys	80
Db	566	CTGTTTCTTTTGAAGTGTGGAAGTGTGACAGTGTGACAGTCAAGAAATTTTGTCTTAAG	625
Qy	81	CysAlaArgAlaGluGluLeuPheAsnMetLeuGlnIleMetGlnAsnSerIle	100

Db	626	TGTGCCCGTGCAAGAAATTATTTAAACATGTTGCAAGAGATTATCCAAAATAAGTATA	685
Qy	101	AsnValValGluGluProValValGluArgAsnHisGlnThrGluLeuGluValPro	120
Db	686	AATGTGGTGAAGACCGCGTTGTAGAAAGAGTAGTATCATCAGACAGACTGGAAGTCCCT	745
Qy	121	ArgThrProArgThrProThrThrProGly	130
Db	746	AGGACACCTCGACACCTACACTCCAGGG	775

RESULT 4

BC055334

LOCUS

DEFINITION

Mus musculus fibroblast growth factor receptor substrate 2, mRNA

ACCESSION

BC055334

VERSION

BC055334.1

KEYWORDS

MGC.

SOURCE

Mus musculus

ORGANISM

Mus musculus (house mouse)

REFERENCE

1. (bases 1 to 5792)

AUTHORS

Strausberg, R.D., Feingold, B.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carrinci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullany, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S., Sanchez, A., Whitting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouford, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butlerfield, Y.S., Krzywinski, M.I., Skalska, U., Smalish, D.E., Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.

Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

12477932

2. (bases 1 to 5792)

AUTHORS

Strausberg, R.

Direct Submission

Submitted (22-JUL-2003) National Institutes of Health, Mammalian

Gene Collection (MGC), Cancer Genomics Office, National Cancer

Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>Contact: cgabbs@mail.nih.gov

Tissue Procurement: Dr. Jim Lin, University of Iowa

cDNA Library Preparation: M. Bento Soares, University of Iowa

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: University of Iowa, Dr. M. Bento Soares and Dr. Thomas L. Casavant.

Web site: <http://genome.uiowa.edu>Contact: bento-soares@uiowa.edu; tom-casavant@uiowa.edu

Ronald, M.F., Akabogu, I., Bair, T., Bair, J., Crouch, K., Davis, A., Fishler, K., Koppel, C., Kucaba, T., Lebeck, M., Melo, A., Schaefer, K., Scheetz, T., Smith, C., Snit, E., Tack, D., Trout, K., Walters, J., Casavant, T., Soares, M.B.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Series: Plate: Row: Column: 0

This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis.

```

FEATURES
  source
    Location/Qualifiers
      1..5792
        /organism="Mus musculus"
        /mol_type="mRNA"
        /strain="C57BL/6"
        /db_xref="taxon:10090"
        /clone="MGC:64739 IMAGE:6830555"
        /tissue_type="Brain, enriched mouse brain 12.5dp"
        /clone_lib="NIH_BMAP_F00"
        /lab_host="DH10B"
        /notes="Vector: pYX-ASC"
      1..5792
        /gene="Frs2"
        /notes="synonyms: 4732458E18, SNT1"
        /db_xref="LocusID:327826"
        /db_xref="MGI:1100860"
      392..1918
        /codon_start=1
        /product="Frs2 protein"
        /protein_id="AAH55334.1"
        /db_xref="GI:33244013"
        /db_xref="LocusID:327826"
        /translation="MGSCCPCPKDTVPDHNKFKVINVDGNGELGSGVMELDTDE
        LILYTKRSVSKWHYICLRRYGYDNLFSFESGRRCQTGGIFAFKARAELEFNMLQ
        EIMQNNISVVEEPVVERSHQTELEVPRTPTPTTGLGAQLNPGYRPSFGDAS
        SHPSRHSVGSARLPSVGBESTHPLLVPEDHVHTVNTGVOEKKNRASHVDPPEA
        RVGNASNTPKRPSNPEDRDPOVLKPGVFLGPTPVOKLMKEKLEGLQKDPV
        SSGAGNTKEDTGYDSDERDVPVNVKLAVENINGLSIPASGVREGLTSTSDTQ
        NINNSQRPALNLYNLSPLPWEARKLRDEDDNLGKTPSLNGYHNNDPHENY
        VNTENVTPLSAHKVFEFRRDCSPTVFNFDIRRPSLEHQRLNYIOVDLGGSDSDN
        QTPTKPTPTTP.PQTPTRTTELYAVIDIERTAMGNLQALPRDDGTSRKTRHNSDLP
        "
      misc_feature
        443..721
          /notes="IRS: Region: PTB domain (IRS-1 type)"
          /db_xref="CDD:pfam02174"

ORIGIN
Alignment Scores:
Pred. No.:      2,93e-77      Length:      5792
Score:          685.00      Matches:      127
Percent Similarity: 100.00%      Conservative: 3
Best Local Similarity: 97.69%      Mismatches: 0
Query Match:      98.42%      Indels: 0
DB:              10          Gaps: 0

US-09-757-415A-1_COPY_11_140 (1-130) x BC055334 (1-5792)
QY 1 AspThrValProAspAsnHisArcAsnLysPheLysValIleAsnValAspAspGly 20
Db 422 GACACTCCAGATACCAACACGACGAAAGTTTAAGTCAATTAATGTGGATGATGGG 481
QY 21 AsnGluLeuGlySerGlyIleMetGluLeuThrAspThrGluLeuIleLeuYrThrArg 40
Db 482 AATGAGCTAGCTCTGGTGTGATGGAACTCACACACACAGAGCTATTCTGTACACCGG 541
QY 41 LysArgAspSerValLysTrpHisTyrLeuCysLeuArgArgTyrGlyTyrAspSerAsn 60
Db 542 AAACGGGACTCGGTAAATAGGCACTACCTCTGCTACGACGATACGCGTATGACTCAAT 601
QY 61 LeuPheSerPheGluSerGlyArgArgCysGlnThrGlyGlnGlyIlePheAlaPheLys 80
Db 602 CTGTTTCTTTTGAAGTGGTCGAGAGGTGTGCAGACTGGACAGGAATTTTGTCTTTAAG 661
QY 81 CysAlaArgAlaGluGluLeuPheAsnMetLeuGlnGluIleMetGlnAsnSerIle 100
Db 662 TGTCCCGCTGCAGAGAATTAATTAACTGTCAGAGATTATGCAAAATAATAGTATA 721
QY 101 AsnValValGluGluProValValGluArgAsnHisGlnThrGluLeuGluValPro 120
Db 722 AATGTGTGGAAGACCGGTTGTAGAAAGGAGTAGTCATCAGACAGAGCTGGAGTCCCT 781
QY 121 ArgThrProArgThrProThrThrProGly 130

```

```

Db 782 AGGACACCTCGGACACACCTACCACTCCAGG 811

RESULT 5
AB064525
LOCUS
DEFINITION
Xenopus laevis xFRS2 mRNA for FRs2, complete cds.
ACCESSION
AB064525
VERSION
AB064525.1 GI:14588669
KEYWORDS
Xenopus laevis (African clawed frog)
SOURCE
Xenopus laevis
ORGANISM
Xenopus laevis
REFERENCE
1 Kusakabe, M., Masuyama, N., Hanafusa, H. and Nishida, E.
Xenopus FRs2 is involved in early embryogenesis in cooperation with
the Src family kinase Laloo
EMBO Rep. 2 (8), 727-735 (2001)
JOURNAL
MEDLINE
21384985
PUBMED
11463744
REFERENCE
2 (bases 1 to 2280)
Kusakabe, M., Masuyama, N. and Nishida, E.
Direct Submission
Submitted (30-JUN-2001) Eisuke Nishida, Graduate school of
Biostudies, Kyoto University, Department of Cell and Developmental
Biology, Kitashirakawa, Sakyo, Kyoto 606-8502, Japan
(E-mail: L50174@sakura.kudpc.kyoto-u.ac.jp, Tel.81-75-753-4230,
Fax:81-75-753-4235)
FEATURES
  source
    Location/Qualifiers
      1..2280
        /organism="Xenopus laevis"
        /mol_type="mRNA"
        /db_xref="taxon:8355"
      1..2280
        /gene="xFRS2"
        /gene="xFRS2"
        /note="involved in early embryogenesis in cooperation with
        the Src family kinase Laloo"
        /codon_start=1
        /product="FRS2"
        /protein_id="BAB61837.1"
        /xref="GI:14588670"
        /translation="MGSCCPCPKDETIPDQNNFKVINVDGNGELGSGMEURENE
        LILYTKRSVSKWPLYCLRRYGYDNLFSFESGRRCQTGGIFAFKARAELEFNMLQ
        EIMQNNISVVEEPVVERNPOTELDVPRTPTPTTGPSSGVNPGYRPSFGAAS
        RPSRHSVGSARLPSVGBESTHPLLVPEDHVHTVNTGVPDQKSRPNAPPAQEVSR
        VSNHEGIDSDAEPVPERDPOVLEPQGVFLGPTPVOROLLEKKNLEKLDQHQASV
        NNDSSCNNTTELDGDSDEEREISSKNKVENLNGLSISSGLRGRGVVPPSTDI
        QVNNASQRRALINENLPSLPWETRKPSREDDSLGKTPSLNGYHNNDPHENY
        VNTENVTPLSAHKVFEFRRDCSPTVFNFDIRRPSLEHQRLNYIOVDLGGSDSDN
        QTPTKPTPTTP.PQTPTRTTELYAVIDIERTAMGNLQALPRDDGTSRKTRHNSDLP
        "
  Alignment Scores:
    Pred. No.:      5,63e-69      Length:      2280
    Score:          616.50      Matches:      117
    Percent Similarity: 95.33%      Conservative: 7
    Best Local Similarity: 90.00%      Mismatches: 5
    Query Match:      88.53%      Indels: 1
    DB:              5          Gaps: 1

US-09-757-415A-1_COPY_11_140 (1-130) x AB064525 (1-2280)
QY 1 AspThrValProAspAsnHisArcAsnLysPheLysValIleAsnValAspAspGly 20
Db 263 GAACTATCCAGATAACCAACACAGATTAAAGTTAATTAAGTTGATGATCATGGT 322
QY 21 AsnGluLeuGlySerGlyIleMetGluLeuThrAspThrGluLeuIleLeuYrThrArg 40
Db 323 AATGAGCTGGCTCTGGCATATGGAATTTGAGAGAAATGAAGAACTATCTTATACACTCGC 382

```

Qy	41	LysArgAspSerValLysTrpHisTyrIleuCysLeuArgArgTyrGlyTyrAspSerAsn	60
Ddb	383	AAGAGAGATTCTGTCAAGTGGCTTACCTCTGCTTGGACGCTATGATGATGATCAAC	442
Qy	61	LeuPheSerPheGluSerGlyArgArgCysGlnThrGlyGlyGlyGlyIlePheAlaPheLys	80
Ddb	443	CTTTTTCCTTTTGAAGTGGAGGAGGTGTGAGATGGACGAAGAATATTTGCTTTTAAG	502
Qy	81	CysAlaArgAlaGluGluLeuPheAsnMetLeuGlnGluIleMetGlnAsnAsnSerIle	100
Ddb	503	TGTGCCCGCGCAGAGAGCTCTTCAATATGCTTCAAGAGATCATGCAAAATAATACATT	562
Qy	101	AsnValValGluGluProValValGluArgAsnAsnHisGlnThrGluLeuGluValPro	120
Ddb	563	AGTGTGTAGAGGAGCTGTAGTTGAACGA---AATCCCGACACCGATGTGGATGTCCA	619
Qy	121	ArgThrProArgThrProThrProGly	130
Ddb	620	AGAACCCCTCGACACCCACTACCCCTGGC	649
RESULT	6		
LOCUS	BC046943	3017 bp	mRNA linear VRT 04-NOV-2003
DEFINITION	Xenopus laevis such1-associated neurotrophic factor target (FGFR signalling adaptor), mRNA (cdna clone MGC:52516 IMAGE:5570479), complete cds.		
ACCESSION	BC046943	GI:28422658	
VERSION	BC046943.1	GI:28422658	
KEYWORDS	MGC.		
SOURCE	Xenopus laevis (African clawed frog)		
ORGANISM	Xenopus laevis		
KEYWORDS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae; Xenopodinae; Xenopus.		
REFERENCE	1 (bases 1 to 3017)		
AUTHORS	Klein,S.L., Strausberg,R.L., Wagner,L., Pontius,J., Clifton,S.W. and Richardson,P.		
TITLE	Genetic and genomic tools for Xenopus research: The NIH Xenopus initiative		
JOURNAL	Dev. Dyn. 225 (4), 384-391 (2002)		
MEDLINE	22341132		
PUBMED	12454917		
REFERENCE	2 (bases 1 to 3017)		
AUTHORS	Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G., Klausner,R.D., Collins,F.S., Wagner,L., Shennan,C.M., Schuler,G.D., Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,P., Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L., Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L., Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S., Carnici,P., Prange,C., Raja,S.S., Loquellano,N.A., Peters,G.J., Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J., McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S., Worley,K.C., Hale,S., Garcia,A.M., Gay,L.D., Hulyk,S.W., Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Fahey,J., Helton,E., Kettman,M., Madan,A., Rodriguez,S., Sanchez,A., Whitting,M., Madan,A., Young,A.C., Shevchenko,K., Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D., Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M., Buttefield,Y.S., Krzywinski,M.I., Skalska,U., Smalusz,D.E., Schnerch,A., Schein,J.B., Jones,S.J. and Marra,M.A.		
TITLE	Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences		
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)		
MEDLINE	22388257		
PUBMED	12477932		
REFERENCE	3 (bases 1 to 3017)		
AUTHORS	Klein,S. and Strausberg,R.		
TITLE	Direct Submission		
JOURNAL	Submitted (14-FEB-2003) National Institutes of Health, Xenopus Gene Collection (XGC), National Institute of Child Health and Human Development, 6100 Executive Boulevard, Room 4B01, Rockville, MD		

Db 204 GAAACTATCCAGATACCAACAAAACAGATTTTAAGGTTATTAAACGTTGATGATGCT 263

Qy 21 AsnGluLeuGlySerGlyIleMetGluLeuThrAspThrGluLeuIleLeuTyrThrArg 40

Db 264 AATGAGCTTGCGCTCTGGCAATAATGGAAATGGAGAAATGAATCACTATCTATACACTCGC 323

Qy 41 LysArgAspSerValIysTrpHisTyrLeuCysLeuArgTyrGlyTyrAspSerAsn 60

Db 324 AAGAGAGATTCTGCTCAAGTGGCTTACCTCTGCTTGGACGCTATGGATACGATTCACAAAC 383

Qy 61 LeuPheSerPheGluSerGlyArgGlyCysGlnThrGlyGlnGlyIlePheAlaPheLys 80

Db 384 CTCTTTCTTTTGAAGTGGACGAGGCTGTCAGACTGGACGAGAAATATTTGCTTTTAAG 443

Qy 81 CysAlaArgAlaGluGluLeuPheAsnMetLeuGlnGlnIleMetGlnAsnAsnSerIle 100

Db 444 TGTGCGCGCGCAGAGAGCTCTTCAATATGCTTCAAGAGATCATGCAAAATAATACGATT 503

Qy 101 AsnValValGluGluProValValGluArgAsnAsnHisGlnThrGluLeuGluValPro 120

Db 504 AGTGTTGTAGAGAGGCTGTAGTTGAACGA---AATCCCGACAGCGAGTTGGATGTCCA 560

Qy 121 ArgThrProArgThrProThrThrProGly 130

Db 561 AGAACCCCTCGACACCCATACCCCTGGC 590

RESULT 7

AF390895 1612 bp mRNA linear VRT 20-FEB-2002

LOCUS

DEFINITION

Xenopus laevis sucl1-associated neurotrophic factor target XSNT

(SNT) mRNA, complete cds.

ACCESSION

AF390895

VERSION

AF390895.1 GI:18033969

KEYWORDS

ORGANISM

Xenopus laevis (African clawed frog)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae; Xenopodinae; Xenopus.

REFERENCE

1 (bases 1 to 1612)

Axagi,K., Kyun Park,E., Mood,K. and Daar,I.O.

Docking protein SNT1 is a critical mediator of fibroblast growth factor signaling during Xenopus embryonic development

Dev. Dyn. 223 (2), 216-228 (2002)

21826201

11836786

REFERENCE

2 (bases 1 to 1612)

Axagi,K., Mood,K. and Daar,I.O.

Direct Submission

Submitted (12-JUN-2001) Regulation of Cell Growth Laboratory, National Cancer Institute, Frederick, MD 21702, USA

LOCATION/Qualifiers

1..1612

/organism="Xenopus laevis"

/mol_type="mRNA"

/db_xref="taxon:8355"

1..1612

/gene="SNT"

1..1530

/gene="SNT"

/note="FGFR signaling adaptor FR2"

/codon_start=1

/product="sucl1-associated neurotrophic factor target XSNT"

/protein_id="AA157304.1"

/db_xref="GI:18033970"

/translation="MGSCCSPDKETIPNQQRFXVINVDDGNELGSGIMELTEND LILTRKDSVRVPLICRRYGDNSLNFSESGRCQTGGGIFAFKCARABELFNMLQ EIMQNTISVVEFPVVERNPOTELDPTPTTGFSGGVNPGTFRYPSPVGEASS HPSRHSPSVGSTRLPYSVEESTHLPVPEHDVHTYNTSGVEDQKQRPVPAQEV VSSPEGHGHEGPPVPRDPQVILEPDGVFLGTPVQRLIEKKLEITEQHDSV NGGSCSNPTNDIYDSDHERTPSNKMVCENLGNLSLSAVQGRVWPPISTDI QNVNNSAQRTALINYLELPSLPVMEATKRPSEEDDSLGKPTPSLNGFHSNLDPMHN YVNTENTVFLSAHKVEFSRRDCTPVFNFDIRRSLEQRQLNYIQVDLEGSDSDN

FEATURES

source

gene

CDS

ORIGIN

Alignment Scores:

Score: 8.65e-69 Length: 1612

613.50 Matches: 115

Percent Similarity: 96.15% Conservative: 10

Best Local Similarity: 88.46% Mismatches: 4

Query Match: 88.15% Indels: 1

DB: 5 Gaps: 1

US-09-757-415a-1_COPY_11_140 (1-130) x AF390895 (1-1612)

Qy 1 AspThrValProAspAsnHisArgAsnLysPheLysValIleAsnValAspAspGly 20

Db 31 GAAACTATCCAGATACCAACAAAACAGATTTTAAGGTTATTAAATGTTGATGATGCT 90

Qy 21 AsnGluLeuGlySerGlyIleMetGluLeuThrAspThrGluLeuIleLeuTyrThrArg 40

Db 91 AATGAGCTTGCGCTCTGGCAATAATGGAAATGGAGAAATGAATCACTATCTATACACTCGC 150

Qy 41 LysArgAspSerValIysTrpHisTyrLeuCysLeuArgTyrGlyTyrAspSerAsn 60

Db 151 AAGCGAGATTCTGTCAGATGGCTTACCTCTGCTTGGACGCTATGGATACGATTCACAAAC 210

Qy 61 LeuPheSerPheGluSerGlyArgGlyCysGlnThrGlyGlnGlyIlePheAlaPheLys 80

Db 211 CTCTTTCTTTTGAAGTGGACGAGGCTGTCAGACTGGACGAGAAATATTTGCTTTTAA 270

Qy 81 CysAlaArgAlaGluGluLeuPheAsnMetLeuGlnGlnIleMetGlnAsnAsnSerIle 100

Db 271 TGTGCGCGCGCAGAGAGCTCTTCAATATGCTTCAAGAGATCATGCAAAATAATACATT 330

Qy 101 AsnValValGluGluProValValGluArgAsnAsnHisGlnThrGluLeuGluValPro 120

Db 331 AGCGTTGTGGAGAGCCAGTAGTTGAGCGA---AATCCCGACAGCTGAGTTGATGTCCA 387

Qy 121 ArgThrProArgThrProThrThrProGly 130

Db 389 AGAACCCCTCGACACCCACCAACCCCTGGA 417

RESULT 8

BD079340 766 bp DNA linear PAT 27-AUG-2002

LOCUS

DEFINITION

Cancer-associated nucleic acids and polypeptides.

ACCESSION

BD079340

VERSION

BD079340.1 GI:22624943

KEYWORDS

JP 2001516009-A/6.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 766)

Old,I.J., Scanlan,M.J., Stockert,B., Gure,A., Chen,Y.T., Gout,I., Oghare,M., Obata,Y., Pfreundschuh,M., Tureci,O. and Sahin,U.

Cancer-associated nucleic acids and polypeptides

Patent: JP 2001516009-A 6 25-SEP-2001;

LUDWIG INSTITUTE FOR CANCER RESEARCH

OS Homo sapiens (human)

PN JP 2001516009-A/6

PD 25-SEP-2001

PF 15-JUL-1998 JP 2000503425

PR 17-JUL-1997 US 08/896164,10-OCT-1997 US 60/061599 PR

10-OCT-1997 US 60/061765,10-OCT-1997 US 08/948705 PR

11-OCT-1997 GB 9721697.2,22-JUN-1998 US 09/102322 PI LLOYD

J OLD,MATTHEW J SCANLAN,ELISABETH STOCKERT,ALI GURE,YAO PI TSENG

CHEN, IVAN GOUT,MICHAEL O'HARE,YUICHI OBATA,MICHAEL PFREUNDSCUH, PI

OZLEM TURECI,

PI UGUR SAHIN

PC

G01N33/574,A61K38/00,A61K39/395,A61K45/00,A61K48/00, PC

POTPKTPTTLPQPTTTRTELYAVIDIERTAAMSNLOKALPRDDGTSRKTRHNSDLP

M"

A61P35/00,
 PC C07K14/82, C07K16/32, C12N15/09//C07K16/46, C12P21/08, A61X37/02,
 PC C12N15/00
 CC Cancer-associated nucleic acids and polypeptides. FH Key
 Location/Qualifiers
 FT source 1..766 /organism="Homo sapiens (human)".
 FT Location/Qualifiers
 1..766
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"

FEATURES

source

ORIGIN

Alignment Scores:
 Pred. No.: 7,43e-57 Length: 766
 Score: 517.50 Matches: 120
 Percent Similarity: 87.77% Conservative: 2
 Best Local Similarity: 86.33% Mismatches: 7
 Query Match: 74.35% Indels: 10
 DB: 6 Gaps: 1

US-09-757-415a-1_COPY_11_140 (1-130) x BD079340 (1-766)

QY 1 AspThrValProAspAsnHisArgAsnLysPheLysValIleAsnValAspAspGly 20
 DB 359 GACACTGTCACAGATAACCATCGGACACAGTTTAAGTCATTATGTTGGATGATGGG 418
 QY 21 AsnGluLeuGlySerGlyIleMetGluLeuThrAspThrGluLeuIleLeuTyrThrArg 40
 DB 419 AATGAGTTAGTTCCTGGCATATGGAACCTACAGACACAGAACTGATTTTATACACCGC 478
 QY 41 LysArgAspSerValLysTyrPheTyrLeuCysLeuArgTyrGlyTyrAspSerAsn 60
 DB 479 AAAGTGAATGCTAGTAATGATGCACTACTCTGCTGGACGCTATGCTATGCTGGAAT 538
 QY 61 LeuPheSerPheGluSerGlyArgGlyCysGlnThr-GlyGlnGlyIlePheAlaPheLeu 80
 DB 539 CTCCTTTCTTTTGAAGTGTGCAAGGTGTCAAACTGGGACAGGAATCTTTGCTTTAA 598
 QY 80 sCysAlaArgAla-GluGluLeuPheAsnMetLeuGln-GluIleMetGln-AsnAsnSe 99
 DB 599 GTGTGCCCGCTGCAAAAGAAATATTATACATGTTGCAAGAGATTATGCAAAATATATAG 658
 QY 99 rIleAsnVal-ValGluGlu-ProVal-ValGluArgAsnAsnHisGln--ThrGluLeu 117
 DB 659 TATAATGTTGTTGGTGAAGACCAATTTGTAGAAGAATAATCATCAACAGCAAGATTG 718
 QY 118 GluValPro---ArgThrProArgThrProThrThrPro 129
 DB 719 GAAGTCCCTTAANAACAACCTCGAACAACATACACTCCA 757

RESULT 9

LOCUS

BC010611 2031 bp mRNA linear PRI 07-OCT-2003
 DEFINITION Homo sapiens fibroblast growth factor receptor substrate 3, mRNA
 (CDNA clone MGC:17167 IMAGE:4183414), complete cds.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

BC010611.1 GI:14714909
 MGC.
 Homo sapiens (human)
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 2031)
 Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
 Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
 Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
 Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, P.,
 Diatchenko, L., Narusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
 Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
 Scheet, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,
 Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,

Abramson, R.D., Mullaby, S.J., Bosak, S.A., McEwan, P.J.,
 McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
 Worley, K.C., Hale, S., Garcia, A.M., Gay, E.J., Hulyk, S.W.,
 Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
 Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodrigues, S.,
 Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
 Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
 Dickinson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
 Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalls, D.E.,
 Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
 Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
 22388257
 MEDLINE
 12477932
 REFERENCE
 2 (bases 1 to 2031)
 Strausberg, R.
 Direct Submission
 Submitted (10-JUL-2001) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 NIH-MGC Project URL: http://mgc.nci.nih.gov
 Contact: MGC help desk
 Email: cgabbs@mail.nih.gov
 Tissue Procurement: David N. Louis, M.D.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILLNL)
 DNA Sequencing by: Baylor College of Medicine Human Genome
 Sequencing Center
 Center code: BCM-HGSC
 Web site: http://www.hgsc.bcm.tmc.edu/cdna/
 Contact: amg@bcm.tmc.edu
 Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Loulseghe, H.,
 Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati,
 A.N., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
 Series: IRAK Plate: 12 Row: 1 Column: 14
 This clone was selected for full length sequencing because it
 passed the following selection criteria: matched mRNA gi: 5730058.

FEATURES

source

1..2031
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="MGC:17167 IMAGE:4183414"
 /tissue_type="Brain, glioblastoma with EGFR amplification"
 /clone_lib="NCI CGAP_Brn64"
 /lab_host="DH10B"
 /note="Vector: pCMV-SPORT6"
 1..2031
 /gene="FRS3"
 /note="synonyms: SNT2, FRS2B, MGC17167, SNT-2, FRS2beta"
 /db_xref="LocusID:10817"
 /db_xref="MIM:507744"
 76..1554
 /codon_start=1
 /product="fibroblast growth factor receptor substrate 3"
 /protein_id="AAH10611.1"
 /db_xref="GI:14714910"
 /db_xref="LocusID:10817"
 /translation="MGCSCCLNRSVDPNHPFKVTNVDDGVEGLSGVMELTQSE
 IVLHLHREAVRWPLYCLRLRYGYSNLFPSFGRRCOTGGIFAFKSRAREINLQ
 DLMQNSINNVMEPVITRNHPALDLPAPQPPNALGYTVSFGSCGEGFRFA
 PRLSTSLRHPGLSEETHALIAIDQSHITYNTVPASDDHRRGRLQLPPEGQAP
 FLCPAQDQDPQVFLQPGQKVLGPTTARRHMKCCGLCPSLHPDPPHNNNEAP
 SECPAQKCTYENTVGLRGCAWRGLSPGPGWNLHARRAALLHVNLPPLPVPWES
 QAQLGGEAGDGDSDGLTPSSNGFPDGETPLOTPTSTRAAISHGSEFPVPLTR
 RSGSFVNFDFRPPGPPFQLNYIQVELKMGWGDGPKGPQNPSPSQAMPPTTPAR
 127..405

CDS

misc_feature

```
/notes="IRS; Region: PTB domain (IRS-1 type)"
/db_xref="CDD:pfam02174"

ORIGIN
Alignment Scores:
Pred. No.: 2,16e-55 Length: 2031
Score: 510.50 Matches: 92
Percent Similarity: 86.26% Conservative: 21
Best Local Similarity: 70.23% Mismatches: 17
Query Match: 73.35% Indels: 1
DB: 9 Gaps: 1

US-09-757-415A-1_COPY_11_140 (1-130) x BC010611 (1-2031)
QY 1 AspThrValProAspAsnHisArgAsnLysPheLysValIleAsnValAspAspGly 20
DB 106 GACAGCGTCCAGACACACACCCACCAGTCAAGGTGACAAATGTGGATGATGAGGG 165
QY 21 AsnGluLeuGlySerGlyIleMetGluLeuThrAspThrGluLeuLeuLeuThrArg 40
DB 166 GTGAGCGTGGGCTCTGGGGGTGATGGAGCTGACGACAGTGGCTGGTGCACCTGCAT 225
QY 41 LysArgAspSerValLysTrpHisTyrLeuCysLeuArgArgTyrGlyTyrAspSerAsn 60
DB 226 CGGGTGAGCGCTCCCGCTGCCTTATCTCTGTTGGGGCGCTATGGCTACGACTCCAC 285
QY 61 LeuPheSerPheGluSerGlyArgCysGlnThrGlyGlnGlyIlePheAlaPheLys 80
DB 286 CTCTTCTCTTTCAGAGTGGCGCGATGTACAGACAGCCAGGAAATATTGCAATTAAG 345
QY 81 CysAlaArgAlaGluLeuPheAsnMetLeuGlnGluIleMetGlnAsnSerIle 100
DB 346 TGTTCCCGGGCTGAGGAATCTTCAACCTCTCAGGATCTGATGAGTGCACACGATC 405
QY 101 AsnValValGluGluProVal---ValGluArgAsnHisGlnThrGluLeuGluVal 119
DB 406 AATGTGTGAAGAGCGTGTGCATCATCACCGGCAATAGCACCCCGCTGAGCTGACCTC 465
QY 120 ProArgThrProArgThrProThrProGly 130
DB 466 CCTCAGCCCCCAGCACCAATGCTCTAGGC 498

RESULT 10
AF036718 2077 bp mRNA linear PRI 30-JUL-1999
LOCUS Homo sapiens FGFR signalling adaptor SNT-2 mRNA, complete cds.
DEFINITION
ACCESSION AF036718
VERSION AF036718.1 GI:2708629
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 2077)
Xu,H., Lee,K.W. and Goldfarb,M.
Novel recognition motif on fibroblast growth factor receptor
mediates direct association and activation of SNT adaptor proteins
J. Biol. Chem. 273 (29), 17987-17990 (1998)
JOURNAL 98324992
MEDLINE
PUBMED 9660748
REFERENCE
2 (bases 1 to 2077)
Xu,H., Lee,K. and Goldfarb,M.
Direct Submission
AUTHORS Submitted (04-DEC-1997) Brookdale Center for Developmental and
Molecular Biology, Mt. Sinai School of Medicine, 1 Gustave Levy
Place, New York, NY 10029, USA
JOURNAL
FEATURES
Location/Qualifiers
1..2077
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue type="placenta"
1..2077
source
gene
```

```
/gene="SNT-2"
137..1615
/gene="SNT-2"
/codon_start=1
/product="FGFR signalling adaptor SNT-2"
/protein_id="AAB32555.1"
/db_xref="GI:2708630"
/translation="MGSCCSCNDSVDPNHPTKFKVTNVDDGVELGSGVNEUTQSE
LVHLHREAARWYPLCLRRYGSNLFSSGRRCTGQGFQFAFKCSRAEIFNLQ
DLMCNSINVMEEPIITRNHSHPAELDIPRAOPNALGYTVSSFGSGCGEGRFSA
PRLSTSLRHSIGESTHALIAPDEQSHYVNTPASDEHRRGRHCLQPLPGQAP
FLPARGPDQDPQVFLQPGOVKVLGPTPARHMKVCOGLCPSLHDPHPHNNKEAP
SECPAQKCTYENVITGLWRGAGRWLSPEFGWNGLAHRAALLHYENLPLPVPWES
CAQLGGAGAGDSDGSDLTSSNGFFDEEDTLPKPTSTRAAIRSHSGFPVPLIR
RRGSPRVNFDPRPEPPQLNYIQLKMGWGDPRKPGQNPSSPOAPMPETHPAR
SSDSYAVTDLKKTVAMSLQRLPRDDGTARKTRNSTDLPL"
137..154
/misc_feature
/gene="SNT-2"
/note="encodes myristoylation motif"
182..532
/misc_feature
/gene="SNT-2"
/note="encodes PTB domain"

ORIGIN
Alignment Scores:
Pred. No.: 2,23e-55 Length: 2077
Score: 510.50 Matches: 92
Percent Similarity: 86.26% Conservative: 21
Best Local Similarity: 70.23% Mismatches: 17
Query Match: 73.35% Indels: 1
DB: 9 Gaps: 1

US-09-757-415A-1_COPY_11_140 (1-130) x AF036718 (1-2077)
QY 1 AspThrValProAspAsnHisArgAsnLysPheLysValIleAsnValAspAspGly 20
DB 167 GACAGCGTCCAGACACACACCCACCAGTTCAGAGTGACAAATGTGGATGAGGG 226
QY 21 AsnGluLeuGlySerGlyIleMetGluLeuThrAspThrGluLeuLeuLeuThrArg 40
DB 227 GTGAGAGCTGGGCTCTGGGGGTGATGGAGCTGACGACAGTGGCTGGTGCACCTGCAT 286
QY 41 LysArgAspSerValLysTrpHisTyrLeuCysLeuArgArgTyrGlyTyrAspSerAsn 60
DB 287 CGGGTGAGCGCTCCCGCTGCCTTATCTCTGTTGGGGCGCTATGGCTACGACTCCAC 346
QY 61 LeuPheSerPheGluSerGlyArgCysGlnThrGlyGlnGlyIlePheAlaPheLys 80
DB 347 CTCTTCTCTTTCAGAGTGGCGCGATGTACAGACAGCCAGGAAATATTGCAATTAAG 406
QY 81 CysAlaArgAlaGluLeuPheAsnMetLeuGlnGluIleMetGlnAsnSerIle 100
DB 407 TGTTCCCGGGCTGAGGAATCTTCAACCTCTCAGGATCTGATGAGTGCACACGATC 466
QY 101 AsnValValGluGluProVal---ValGluArgAsnHisGlnThrGluLeuGluVal 119
DB 467 AATGTGTGAAGAGCGTGTGCATCATCACCGGCAATAGCACCCCGCTGAGCTGACCTC 526
QY 120 ProArgThrProArgThrProThrProGly 130
DB 527 CCTCAGCCCCCAGCACCAATGCTCTAGGC 559

RESULT 11
BC014819 2150 bp mRNA linear ROD 06-OCT-2003
LOCUS Mus musculus fibroblast growth factor receptor substrate 3, mRNA
DEFINITION (CDNA clone MGC:25496 IMAGE:4506982), complete cds.
ACCESSION BC014819
VERSION BC014819.1 GI:15928711
KEYWORDS MGC.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```


Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 2150)
 Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
 Klausner, R.D., Collins, P.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
 Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
 Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Haieh, F.,
 Datchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
 Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
 Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,
 Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
 Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J.,
 McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
 Worley, K.C., Hale, S., Garcia, A.M., Gay, D.J., Hulyk, S.W.,
 Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
 Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S.,
 Sanchez, A., Whitting, M., Madan, A., Young, A.C., Shevchenko, Y.,
 Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
 Dickson, M.C., Rodriguez, A.C., Touchman, J.W., Green, E.D.,
 Butterfield, Y.S., Kzywinski, M.I., Skalska, U., Smalios, D.E.,
 Schnerch, A., Schein, J.E., Jones, S.J., and Marra, M.A.
 Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
 2388257
 12477932
 2 (bases 1 to 2150)
 Strausberg, R.
 Direct Submission
 Submitted (01-OCT-2001) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: The Cepko Laboratory
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)
 DNA Sequencing by: Baylor College of Medicine Human Genome
 Sequencing Center
 Center code: BCM-HGSC
 Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
 Contact: amg@bcm.tmc.edu
 Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Loulseged, H.,
 Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati,
 A.N., Gibbs, R.A.
 Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/ILNL at: <http://image.llnl.gov>
 Series: IRAC Plate: 31 Row: m Column: 15
 This clone was selected for full length sequencing because it
 passed the following selection criteria: matched mRNA gi: 21450352.
 Location/Qualifiers
 1..2150
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 /clone="MGC:25496 IMAGE:4506982"
 /tissue_type="Eye, retina, mouse strain C57Bl/6"
 /clone_lib="NIH_MGC_94"
 /lab_host="DH10B"
 /note="vector: pCW-SPORT6"
 1..2150
 /gene="Fre3"
 /note="synonym: MGC25496"
 /db_xref="LocustID:107971"
 /db_xref="MGI:2135965"
 233..1711
 /codon_start=1
 /product="fibroblast growth factor receptor substrate 3"
 /protein_id="AAH14819.1"
 /db_xref="GI:15928712"
 /db_xref="LocustID:107971"

/translation="MGSCWSCLDSDRVSPHNHPTKFTVNDDEGVLSGWNELTQSE
 LVHLHOREAVRPYLCLRRYQSDNLFSGRRRCQTGQIFAFKCSRAEDIFNLQ
 DLMQNSINVTBEPVITRSSHPPELDLPRGPQOPAGYVSGFNGFCGCEGPRFS
 SAPRRPSTSLRHPSPGSESTHTLIASEQSHYVNTPTGDEGRSRHCLQLEPGRV
 PLPACTQSDORDPOVLQPGQVFLGPTPARQVMKQCSLCPQMOPPPHNNNEGP
 SECQAPQKCTVENYSGGLQAGWRLSPERGWGLAHRRAALLEYENLPLPWVES
 QQQGGAGAGDGDSDRGLTFSSNGFPDEDEDTPLQPTSTRASARSHSGFPVPLTR
 RGSPPVFNFPDPRQGPFPPLNYIQVLELWGTARPKPQNPSPVSGAPGTPHPVR
 SSDSVAVIDLKKTATMSDLQALPRDDGAVRKRHNSTDLPL"
 284..562
 /note="IRS; Region: PTB domain (IRS-1 type)"
 /db_xref="CDD:pfam02174"
 misc_feature
 ORIGIN
 Alignment Scores:
 Pred. No.: 5,2e-52 Length: 2150
 Score: 484.50 Matches: 88
 Percent Similarity: 85.83% Conservative: 21
 Best Local Similarity: 69.23% Mismatches: 17
 Query Match: 69.61% Indels: 1
 DB: 10 Gaps: 1
 US-09-757-415a-1_copy_11_140 (1-130) x BC014819 (1-2150)
 QY 1 AspThrValProAspAsnHisArgAsnLysPheLysValIleAsnValAspAspGly 20
 DB 263 GACAGGTGCCACACACCCCTACCAAGTTCAGGTGACCAATGTGATGATGAGGG 322
 QY 21 AsnGluLeuGlySerGlyIleMetGluLeuThrAspThrGluLeuLeuThrArg 40
 DB 323 GTGAGTGTGGGTGGGGGTGATGAGCTGACCCAGAGTGGTGGTCTACACCTGCAC 382
 QY 41 LysArgAspSerValLysTrpHisTyrLeuCysLysLeuArgArgTyrGlyTyrAspSerAsn 60
 DB 383 CAGCGGAGAGTGTCCCTGGCCCTACCTCTGCGCGCGCTATGCTGACATCTCAAT 442
 QY 61 LeuPheSerPheGluSerGlyArgArgCysGlnThrGlyGlnGlyIlePheAlaPheLys 80
 DB 443 CTCCTCTCTTTGAGAGTGGCGCGCGCTGTCAGACAGCCAGGGCATCTTTGCTTCAAG 502
 QY 81 CysAlaArgAlaGluLeuPheAsnMetLeuGlnGluIleMetGluAsnSerIle 100
 DB 503 TCTTCGAGGCGCAGAGCATCTTCAACCTGCTGCGAGSACCTCATGCTGAGTAAACAGATC 562
 QY 101 AsnValValGluGluProVal---ValGluArgAsnAsnHisGlnThrGluLeuGluVal 119
 DB 563 ACCTGACGCGAGGAGCCCGTCTATCATCAGCGGAGCAGCCACCCACAGAGCTGACCTG 622
 QY 120 ProArgThrProArgThrPro 126
 DB 623 CCTCGAGGCGCCGCCACAGGCC 643
 RESULT 12
 HSU91327
 LOCUS HSU91327 129252 bp DNA linear PRI 10-JAN-2000
 DEFINITION Human chromosome 12p15 BAC clone CIT987SK-99D8 complete sequence.
 ACCESSION U91327
 VERSION U91327.1 GI:1871209
 HG. HTG.
 KEYWORDS Homo sapiens (human)
 SOURCE
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 129252)
 LOFTUS, B.J., Kim, U.J., Sneddon, V.P., Kalush, F., Brandon, R.,
 Fuhrmann, J., Mason, T., Crosby, M.L., Barnstead, M., Cronin, L.,
 Deslattes Mays, A., Cao, Y., Xu, R.X., Kang, H.L., Mitchell, S.,
 Eichler, E.E., Harris, P.C., Venter, J.C. and Adams, M.D.
 Genome duplications and other features in 12 Mb of DNA sequence
 from human chromosome 16p and 16q
 Genomics 60 (3), 295-308 (1999)
 JOURNAL 99425270
 MEDLINE
 PUBMED 10493829

REFERENCE	2. (bases 1 to 129252)	repeat_region	14625. .14650
AUTHORS	Adams,M.D., Loftus,B.J., Zhou,L., Phillips,C., Brandon,R., Fuhrmann,J., Kim,U.J., Kervage,A.R. and Venter,J.C.	repeat_region	/rpt_family="AT_rich"
TITLE	Chromosome 16p12 BAC Clone C1997SK-99D8 complete sequence	repeat_region	/rpt_family="AluY"
REFERENCE	3. (bases 1 to 129252)	repeat_region	complement(15115. .15156)
AUTHORS	Adams,M.D., Loftus,B.J., Phillips,C.A., Zhou,L., Brandon,R. and Venter,J.C.	repeat_region	/rpt_family="(TAAA)n"
TITLE	Direct Submission	repeat_region	complement(15157. .15455)
JOURNAL	Submitted (04-MAR-1997) The Institute for Genomic Research, 9712 Medical Center Dr., Rockville, MD 20850, USA	repeat_region	/rpt_family="AluY"
REFERENCE	4. (bases 1 to 129252)	repeat_region	complement(15628. .15691)
AUTHORS	Adams,M.D., Loftus,B.J., Zhou,L., La Bombard,M., Kim,U.J. and Venter,J.C.	repeat_region	/rpt_family="AT_rich"
TITLE	Direct Submission	repeat_region	15702. .15887
JOURNAL	Submitted (21-AUG-1997) The Institute for Genomic Research, 9712 Medical Center Dr., Rockville, MD 20850, USA	repeat_region	/rpt_family="MER3"
COMMENT	Genes were identified by a combination of five methods: XGRAIL (available by anonymous ftp from arthur.epm.crn1.gov), Genefinder (available by anonymous ftp from colin@u.washington.edu), GENSCAN (available using the e-mail server at genscan@genomic.stanford.edu), searches of the EST database at TIGR (http://www.tigr.org/cdb/hcd/hcd.html) and searches against a peptide database. Repeats were identified using RepeatMasker (Snit, A.F.A. and Green, P. unpublished, http://ftp.genome.washington.edu/rm/RepeatMasker.html).	repeat_region	15958. .15999
FEATURES	Location/Qualifiers	repeat_region	/rpt_family="MIR2"
source	1. 129252	repeat_region	complement(16139. .16169)
	/organism="Homo sapiens"	repeat_region	/rpt_family="(CA)n"
	/mol_type="genomic DNA"	repeat_region	complement(16304. .16426)
	/db_xref="taxon:9606"	repeat_region	/rpt_family="(TAAAA)n"
	/chromosome="12"	repeat_region	complement(16802. .17097)
	/map="12p15"	repeat_region	/rpt_family="AluSc"
	/clone="C1997SK-99D8"	repeat_region	complement(17507. .17847)
repeat_region	32. .343	repeat_region	/rpt_family="L1MB2"
repeat_region	/rpt_family="AluSq"	repeat_region	complement(17868. .18033)
repeat_region	694. .809	repeat_region	/rpt_family="FAM"
repeat_region	1064. 1347	repeat_region	complement(18434. .18673)
repeat_region	/rpt_family="AluJb"	repeat_region	/rpt_family="MER42c"
repeat_region	complement(1568. .1795)	repeat_region	complement(18686. .18990)
repeat_region	/rpt_family="MIR"	repeat_region	/rpt_family="AluJo"
repeat_region	complement(1989. .2095)	repeat_region	complement(19063. .19362)
repeat_region	/rpt_family="MIR"	repeat_region	/rpt_family="MER42b"
repeat_region	2759. .2842	repeat_region	complement(20433. .20610)
repeat_region	/rpt_family="MIR"	repeat_region	complement(20586. .20757)
repeat_region	complement(2868. .2894)	repeat_region	/rpt_family="L1MD2"
repeat_region	/rpt_family="AT_rich"	repeat_region	complement(20651. .21105)
repeat_region	3070. .3369	repeat_region	complement(20755. .21098)
repeat_region	/rpt_family="AluSx"	repeat_region	complement(20781. .20846)
repeat_region	4939. .5090	repeat_region	/rpt_family="L1MD1"
repeat_region	complement(5370. .5538)	repeat_region	complement(21091. .21112)
repeat_region	/rpt_family="MIR"	repeat_region	/rpt_family="AT_rich"
repeat_region	5532. .5618	repeat_region	complement(21500. .21799)
repeat_region	complement(6341. .6648)	repeat_region	/rpt_family="AluY"
repeat_region	/rpt_family="AluYb8"	repeat_region	22381. .22499
repeat_region	7061. .7195	repeat_region	/rpt_family="FLAM_A"
repeat_region	/rpt_family="MIR"	repeat_region	22412. .22676
repeat_region	7275. .7317	repeat_region	22508. .22552
repeat_region	complement(9247. .9544)	repeat_region	/rpt_family="(CA)n"
repeat_region	/rpt_family="AluJo"	repeat_region	complement(22591. .22891)
repeat_region	complement(10352. .10457)	repeat_region	/rpt_family="AluY"
repeat_region	/rpt_family="MIR"	repeat_region	complement(22894. .23177)
repeat_region	complement(11005. .11292)	repeat_region	/rpt_family="AluSq"
repeat_region	/rpt_family="AluSc"	repeat_region	complement(23341. .23367)
repeat_region	complement(14324. .14349)	repeat_region	/rpt_family="AT_rich"
repeat_region	/rpt_family="AT_rich"	repeat_region	23553. .26023
repeat_region	complement(14400. .14535)	repeat_region	/rpt_family="TIGGER1"
repeat_region	/rpt_family="AluY"	repeat_region	complement(26443. .26597)
		repeat_region	/rpt_family="MIR"
		repeat_region	26810. .27100
		repeat_region	/rpt_family="AluSx"
		repeat_region	complement(28511. .28808)
		repeat_region	/rpt_family="AluJb"
		repeat_region	complement(28922. .29193)
		repeat_region	/rpt_family="AluJo"
		repeat_region	29434. .29472
		repeat_region	/rpt_family="AT_rich"
		repeat_region	30851. .31152
		repeat_region	/rpt_family="AluSp"
		repeat_region	complement(31728. .31914)
		repeat_region	/rpt_family="MIR"
		repeat_region	complement(31932. .32039)
		repeat_region	/rpt_family="MER20"

Elhaj, C., Emerling, S., Escotto, M., Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P., Garrel, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R., Gissel, J. H., Guevara, W., Gunatratne, P., Hale, S., Hamilton, K., Han, J., Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., Hernandez, J., Hernandez, O., Hodgson, A., Hogue, M., Holloway, C., Hollins, B., Honsi, F., Howard, S., Huber, J., Hulyk, S., Hume, J., Ioshikhes, I., Jackson, L. E., Jacobson, B., Jia, X., Johnson, R., Jolivet, S., Joudan, S., Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C., Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lee, E., Lewis, L. C., Lewis, L., Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W., Loulseghe, H., Lozano, R. J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J., Maheshwari, M., Mapua, P., Marondel, I., Martin, R., Martindale, A., Martinez, E., Massey, E., Mawhney, E., McLeod, M. P., Meador, M., Mei, G., Merscher, S., Metzger, M., Miller, A., Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Montgomery, K. T., Morgan, M., Morris, S., Moser, M., Neal, D., Nelson, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N., Nguyen, N., Nickerson, E., Nwokenwo, S., Ogih, M., Okwuonu, G., Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L., Peters, L., Pickens, R., Primus, E., Pu, L. L., Quiles, M., Ren, Y., Rives, M., Rojas, A., Rojokan, I., Rolfe, M., Ruiz, S., Savery, G., Scherer, S., Scott, G., Shen, H., Shim, C., Shoostari, N., Sisson, I., Sodergren, E., Sonalke, T., Sparks, A., Stanley, H., Stone, H., Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H., Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S., Usmani, K., Vasquez, L., Vera, V., Villalon, D., Vinson, R., Wall, R., Wang, S., Ward-Moore, S., Warren, R., Washington, C., Wallington, S., Williams, G., Williamson, A., Wlarczyk, R., Wooden, S., Worley, K., Wu, C., Wu, Y. F., Zhou, J., Zorrilla, S., Zuchel, R., Weinstock, G. and Gibbs, R.

Direct Submission
Unpublished
2 (bases 1 to 188666)
Worley, K.C.

Direct Submission
Submitted (23-DEC-1999) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 188666)
Worley, K.C.

Direct Submission
Submitted (03-NOV-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
4 (bases 1 to 188666)
Worley, K.C.

Direct Submission
Submitted (23-APR-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
5 (bases 1 to 188666)
Worley, K.C.

Direct Submission
Submitted (31-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On Oct 23, 2001 this sequence version replaced gi:16304255.
INFORMATION: <http://www.hgsc.bcm.tmc.edu/> or email gc-help@bcm.tmc.edu

CLONE LENGTH: This sequence does not necessarily represent the entire insert of this clone. Overlapping regions of clones are only sequenced and submitted once, so the sequence for the remainder of the insert may be found in the record for the adjacent clones. Overlapping clones are noted at the beginning and end of the Features listing.

ANNOTATION OF FEATURES:

STSs are identified using ePCR (Genome Res. 7:541-550) searches of a local database that includes entries from dbSTS, GDB, and local mapping efforts.
Repeats are identified using RepeatMasker (A. Smit and P. Green, unpublished.) for Human and Mouse sequences.

Genes and Region of sequence similarity are identified by BLAST (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the EST and cDNA sequences. Genes demonstrate at least two exons flanked by consensus splice sites that maintained sequence continuity across the splice junctions. Sequences that are not identical matches are annotated as similar.

SEQUENCING READ COVERAGE: Sequencing is completed to a minimum standard of double strand coverage with a minimum of 2 clones and 2 reads with no ambiguities or 2 chemistries with a minimum of 2 clones and 3 reads with no ambiguities. If the sequence quality for a region does not meet this standard, it will be indicated in the annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality standards - estimated error rate less than 1 per 10,000 bases. Reports of lowest quality individual bases and measures of base quality are listed below. Description of the metrics can be found at URL:
<http://gc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html>.

FEATURES	Location/Qualifiers
source	1..188666
	/organism="Homo sapiens"
	/mol_type="genomic DNA"
	/db_xref="taxon:9606"
	/chromosome="12"
	/clone="RP11-956E11"
misc_feature	1..2095
	/note="overlaps bases 8907..11001 of clone AC016153"
STS	83..265
	/function="clone overlap"
STS	634..814
	/standard_name="163998"
repeat_region	/standard_name="73530"
	complement(722..786)
	/rpt_family="MER102"
repeat_region	complement(1439..1715)
	/rpt_family="AluX"
repeat_region	2301..2606
	/rpt_family="AluJc"
repeat_region	2884..2908
	/rpt_family="(TTTG)n"
repeat_region	complement(2930..3037)
	/rpt_family="FLAM_A"
repeat_region	3065..3111
	/rpt_family="AT-rich"
repeat_region	3231..3536
	/rpt_family="AluX"
repeat_region	3540..3682
	/rpt_family="MIR"
repeat_region	3761..3999
	/rpt_family="MIR"
repeat_region	4921..5026
	/rpt_family="AluSp"
repeat_region	5515..5550
	/rpt_family="(CA)n"
STS	5786..5959
	/standard_name="52485"
repeat_region	5813..5890
	/rpt_family="L2"
repeat_region	6008..6235
	/rpt_family="MLT1"
repeat_region	6521..6762
	/rpt_family="AluX"
repeat_region	complement(6969..7020)
	/rpt_family="MIR"
repeat_region	7021..7321
	/rpt_family="AluY"
repeat_region	complement(7322..7407)
	/rpt_family="MIR"
repeat_region	8026..8584

```

/rpt family="L2"
8937_9247
/rpt family="AlusX"
9975_10084
/rpt family="MIR"
10159_10228
/rpt family="MLT2G"
complement(10511..10580)
/rpt family="MIR"
complement(10710..11055)
/rpt family="MLT1A1"
11103_11400
/rpt family="AlusG"
11401_11420
/rpt family="(GA)n"
11421_11725
/rpt family="AluY"
11726_11905
/rpt family="(GGAA)n"
complement(12127..12502)
/rpt family="L2"
12530_12792
/rpt family="MER4A2"
12793_13040
/rpt family="AlusX"
13041_13348
/rpt family="MER4A2"
complement(13806..14790)
/rpt family="L2"
complement(15218..15368)
/rpt family="MIR"
15606_15641

Alignment Scores:
Pred. No.: 3.03e-45 Length: 188666
Score: 452.00 Matches: 115
Percent Similarity: 31.27% Conservative: 1
Best Local Similarity: 31.00% Mismatches: 254
Query Match: 64.94% Indels: 1
DB: 9 Gaps: 1

US-09-757-415A-1_COPY_11_140 (1-130) x AC018921 (1-188666)

QY 11 PhelysValIleAsnValAspAspAspGlyAsnGluLeuGlySerGlyIleMetGluLeu 30
Db 73555 TTTTAGTCAATTAATGGATGATGGGAATGAGTAGGTTCTGGCATATGGAACTT 73496

QY 31 ThrAspThrGluLeuIleLeuTyrThrArgLysArgAspSerValIlystrpHisTyrLeu 50
Db 73495 ACAGACACAGAACTGATTTTATACCCGCAACGGTACTCAGTAAATGGCACTACCTC 73436

QY 51 CysLeuArgArgTyrGlyTyrAspSerAsnLeuPheSerPheGluSerGlyArgArgCys 70
Db 73435 TGCGTGGAGCTATGCTATGACTCGATCTCTTTCTTTTGAAGTGGTGAAGGTGT 73376

QY 71 GlnThrGlyGlnGlyIle----- 76
Db 73375 CAACCTGGACAGGTAG-AACCTTGTGTTTTTCCCAATATGTTATTTGAAATGTTT 73317

QY 76 ----- 76
Db 73316 TTTTCAGCTATTCTGTATACAAAGATTAAATTAATTTTCTTCTGAAAAAATCCCAA 73257

QY 76 ----- 76
Db 73256 ATAAACAGATTGTTAAGGAAGATAGTATAGTATTAATTTATCACTGTTTAT 73197

QY 76 ----- 76
Db 73196 AAGCACATTGCTATGTGAGAAATTTTGGCTCATGTCACCTTAAGATTTCAGTAAAT 73137

QY 76 ----- 76

```

```

Db 73136 CATGGGTGACGTAAATTAAGTGTCTTTATCATTTGATTTTTTACTACTACTACT 73077
QY 76 ----- 76
Db 73076 TGAATGGCTTGACAAAGTTACAGGACTCCATTGAAATGTTAAATTTCTTCTATAAAG 73017
QY 76 ----- 76
Db 73016 GGAAATGGTAATACCTATCTTTTAATATTGTTAATAATTAACAAGATAACAATTGAAA 72957
QY 76 ----- 76
Db 72956 ACACCTAGGACTTTCCAAAGGGGCTGTTCATCGCAAAATGATCACCTTGAATAGTAG 72897
QY 76 ----- 76
Db 72896 TTGGGCTCCCTGCTGTGTTAAGAAACAGACTAGCGTAAGTGGCAGGTCCTTGATTTCATT 72837
QY 76 ----- 76
Db 72836 AGCTTTATTTGGGCCATCTCCCTTAGATGAATGAAAAATTTCTGTGAACCTTTGGCGG 72777
QY 76 ----- 76
Db 72776 GGGTTTGTGTTGAACACAGAGATAATCATTAATTTATCAGATATTATTTAAAGGGATT 72717
QY 76 ----- 76
Db 72716 TCTGACTCGGTGCGTAGAATGCTTACTTTACAATTCGTCTGTTGTCAAGTGTTCCTATA 72657
QY 77 -----Phe 77
Db 72656 GTTTTTTTAAAGGTATGTTAACTATTTTCCCTTTGGTTTATTTTGTAGGAATCTTT 72597
QY 78 AlaPheLysCysAlaArgAlaGluGluLeuPheAsnMetLeuGlnGluIleMetGlnAsn 97
Db 72596 GCTTTTAAGTGTGCGGTGCGAGAGAAATTTTAACTGTTGCAAGAGATTATGCAAAAT 72537
QY 98 AsnSerIleAsnValValGluProValValGluArgAsnAsnHisGlnThrGluLeu 117
Db 72536 AATAGTATAAATGTGGTGAAGCCAGTGTGTGAAGAAATAATCATCATCAGAGAATTG 72477
QY 118 GluValProArgThrProArgThrProThrThr 128
Db 72476 GAAGTCCCTAGAACACCTCGAACACCTACAAGT 72444

RESULT 14
AL589661
LOCUS
DEFINITION
  Mouse DNA sequence from clone RP23-58B7 on chromosome 15, complete
  sequence.
ACCESSION
  AL589661
VERSION
  AL589661.21 GI:18476659
KEYWORDS
  HTG.
SOURCE
  Mus musculus (house mouse)
ORGANISM
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
  Dunn,M.
REFERENCE
  1 (bases 1 to 241432)
AUTHORS
  Direct Submission
TITLE
  Submitted (30-JAN-2002) Wellcome Trust Sanger Institute, Hinxton,
  Cambridgeshire, CB10 1SA, UK. E-mail: enquiries@
  hunquerry@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
  On Feb 1, 2002 this sequence version replaced gi:18151481.
  During sequence assembly data is compared from overlapping clones.
  Where differences are found these are annotated as variations
  together with a note of the overlapping clone name. Note that the
  variation annotation may not be found in the sequence submission
  corresponding to the overlapping clone, as we submit sequences with
  only a small overlap as described above.
  This sequence was finished as follows unless otherwise noted: all
  regions were either double-stranded or sequenced with an alternate

```

chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em., EMBL; Sw., SWISSPROT; Tr., TREMBL; Wp., WORMPEP; Information on the WORMPEP database can be found at

http://www.sanger.ac.uk/Projects/C_elegans/wormpep RP23-58B7 is from the RP23-58B7 Mouse PAC Library

constructed by the group of Pieter de Jong.

For further details see <http://www.chori.org/bacpac/home.htm>

VECTOR: pBACE3.6

This sequence is the entire insert of clone RP23-58B7.

FEATURES

source

1. 241432
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosomes="15"
/clone="RP23-58B7"
/clone_lib="RP23-58B7"

misc_feature

147938..148190
/note="Sequence from uni-directional dGTP big dye terminator reads only."

ORIGIN

Alignment Scores:

Prod. No.: 1-25e-39 Length: 241432
Score: 409.50 Matches: 113
Percent Similarity: 24.32% Conservative: 4
Best Local Similarity: 23.49% Mismatches: 1
Query Match: 58.84% Indels: 363
DB: 10 Gaps: 1

US-09-757-415a-1_COPY_11_140 (1-130) x AL589661 (1-241432)

QY 11 PheLysValIleAsnValAspAspGlyAsnGluLeuGlySerGlyIleMetGluLeu 30
|||
Db 95046 TTCTAGTCATTAAATGGATGATGATGGGAATGAGCTAGGCTGTGGTGTGTAAGCACTC 95105
|||
QY 31 ThrAspThrGluLeuIleLeuTyThrArgLysArgAspSerValLysTrpHisTyIleu 50
|||
Db 95106 ACAGACACAGAGCTGATTCGTACACCCGGGAACGGGACTCGGTAAATGGCACTACCTC 95165
|||
QY 51 CysLeuArgArgTyGlyTyAspSerAsnLeuPheSerPheGluSerGlyArgArgCys 70
|||
Db 95166 TGCCTACGACGATACGGCTATGACTCAAAATCTGTTTCTTTTGAAGTGTGCGAAGGTGT 95225
|||
QY 71 GlnThrGlyGln----- 74
|||
Db 95226 CAGACTGGACAGGTAGGGAGAGCCCTTTCTTTCTCCAGATGATAGTGTGGAGCAATT 95285
|||
QY 74 ----- 74
|||
Db 95286 GCACCTCCAGTTACTCTGAATACACATGCGCATGGAGACACAGGCTTGGCCCGAGCC 95345
|||
QY 74 ----- 74
|||
Db 95346 TCGGTGGAGGTAGACATAAGTTATGGATTCAATTACCAAGCATGCGACGACACTGTGCC 95405
|||
QY 74 ----- 74
|||
Db 95406 CTTGATAGAAATAAAGTGTAACTGTCTATGTCCAGCAGTGGAGCTGTGCAGTAGCTAACA 95465
|||
QY 74 ----- 74
|||
Db 95466 GCTTTAACAATCCTGTTCCACAGCAGCAGGAGGCGGTGGCTGGCTGGCTTCCC 95525
|||
QY 74 ----- 74
|||
Db 95526 AGTGACACATGGCGCTCACACACACAGGCCACACCTCTCTAGTGTCTCATAGGAACCTAGA 95585
|||

QY 74 ----- 74
|||
Db 95586 TTGAGATATATAGTCTCTGGGAGCCATTGAGACGCCCGGTCCTGTTTCAATGTC 95645
|||
QY 74 ----- 74
|||
Db 95646 TTCCTTTCTGAGTGTCTTCTCACACTTAGAAAGTCTCAATACATAGATCCATTATACAT 95705
|||
QY 74 ----- 74
|||
Db 95706 AAAGTGGCTTTANTGACATTTTTCACAAATGTACATAATAAAGTGTGATCACACCCACCT 95765
|||
QY 74 ----- 74
|||
Db 95766 CTCACCTCTTGCTAATCTTGTAGCTAGATAGTCCGCGCCCCCCCCCCCCCAATCTAC 95825
|||
QY 74 ----- 74
|||
Db 95826 TTTACATCATTTGGTGTGACGGGCTAGAGGGATGACAGCTTGGGAAATTTGAGTGGGT 95885
|||
QY 74 ----- 74
|||
Db 95886 TTCATTAGAATCATTTACGGGGACATGTGTGAGGGTTTGTTCAGGGGTATAGGCACCTT 95945
|||
QY 74 ----- 74
|||
Db 95946 ACCAAAGGCTGCCCCCATTTGATAAGTATATATCTTCCAAAGCCCCCTTGAGCCGCTCTCT 96005
|||
QY 74 ----- 74
|||
Db 96006 ACCCAAGGTAGGTATGTTGATGGTCTAATAAAGTTCAATGTCATGTCACGGTAACA 96065
|||
QY 74 ----- 74
|||
Db 96066 GCTGTGCGATGCCAGAAATCAGCAATTCAGACCACTCTCTTGCCCTTTCTGTGAGTCTG 96125
|||
QY 74 ----- 74
|||
Db 96126 ACAGTTTCTGCCCTTCTGTACTGTTCTGAGCGTTAGAGGGGAAATACGACCAATGT 96185
|||
QY 74 ----- 74
|||
Db 96186 TCAACAGTCAATTTGTTTTTGGTCTTTGAATCAATCTCTGCAGTTATTGCTGCCTAATGC 96245
|||
QY 74 ----- 74
|||
Db 96246 AAAAGGAAATCTTTTGACCAAGTTCCTATAGCTTGTAAAGGCATGGTTACCATGTTT 96305
|||
QY 75 ----- GlyIlePheAlaPheLysCysAlaArgAlaGluLeu 87
|||
Db 96306 TGCTCTTGATTTATATTTTAGGAATTTTCTTTTAAAGTGTGCCGTGCAGAGAAATTA 96365
|||
QY 88 PheAsnMetLeuGlnGluIleMetGlnAsnAsnSerIleAsnValValGluGluProVal 107
|||
Db 96366 TTTAAATGTTGCAAGAGATTATGCAAAATAATAGTATAATAGTGTGGTGAAGAGCCGGTT 96425
|||
QY 108 ValGluArgAsnAsnHisGlnThrGluLeuGluValProArgThrProArgThrProThr 127
|||
Db 96426 GTAGAAAGGATGATGTCATCAGACAGAGCTGGAAGTCCCTAGACACCTTCGACACCTACA 96485
|||
QY 128 Thr 128
|||
Db 96486 AGT 96488
|||

RESULT 15

AC106185

LOCUS

DEFINITION

Rattus norvegicus clone CH230-141M9, WORKING DRAFT SEQUENCE, 3 unordered pieces.

ACCESSION

AC106185.4 GI:30579183

VERSION
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.

AC106185 238379 bp DNA linear HTG 13-MAY-2003

in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
Center: Baylor College of Medicine
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center Project name: GUP
Center Clone name: CH230-141M9
----- Summary Statistics
Assembly program: Atlas 3.0;
Consensus quality: 232564 bases at least Q40
Consensus quality: 234634 bases at least Q30
Consensus quality: 236042 bases at least Q20
Estimated insert size: 246019; sum-of-contigs estimation
Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 3 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 235027: contig of 235027 bp in length
* 235028 235127: gap of unknown length
* 235128 236230: contig of 1103 bp in length
* 236231 236330: gap of unknown length
* 236331 238379: contig of 2049 bp in length.
----- Location/Qualifiers
1..238379
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-141M9"
1..2257
/note="wgs end_extension
clone_end:T7"
3173..4047
/note="clone_boundary
clone_end:T7"
site:EcoRI
end_sequence:BZ094321"
3203..4009
/note="clone_boundary
clone_end:T7"
site:EcoRI
end_sequence:BZ094321"

FEATURES
source

misc_feature

misc_feature

misc_feature

ORIGIN

Alignment Scores:

Prod. No.: 1-27e-30 Length: 238379
Score: 339.50 Matches: 73
Percent Similarity: 62.69% Conservative: 11
Best Local Similarity: 54.48% Mismatches: 31
Query Match: 48.78% Indels: 19
DB: 2 Gaps: 3

US-09-757-415A-1_COPY_11_140 (1-130) x AC106185 (1-238379)

QY

13 ValletAenValapAppAspGlyAsnGluLeuclySerGlyfileMerGluLeuThasp 32

SOURCE
ORGANISM
Rattus norvegicus (Norway rat)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

REFERENCE
AUTHORS
1 (bases 1 to 238379)
Muzny, D. Marie., Metzker, M. Lee., Abramson, S., Adams, C., Alder, J.,
Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D.,
Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H.,
Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,
Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,
Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,
Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A.,
Chacko, J., Chavez, B., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,
Cleveland, C., Cockrell, R., Cox, C., Coyle, W., Cree, A., D'Souza, L.,
Devila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,
Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,
Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,
Egan, A., Escotto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G.,
Fernandez, S., Finley, M., Flegg, N., Forbes, L., Foster, P.,
Fraser, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,
Guregoris, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W.,
Gunaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K.,
Harvey, Y., Havlak, P., Haves, A., Henderson, N., Hernandez, J.,
Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hogues, M.,
Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A.,
Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,
Karpach, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C.,
Kowis, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, D., Li, Z., Liu, J.,
Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,
Lorenshuwa, L., Louisgied, H., Lozada, R. J., Lu, X., Ma, J.,
Maheshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A.,
Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E.,
Mawhney, S., McLeod, M. P., McNeill, T. Z., Meenen, E.,
Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,
Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L.,
Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S.,
Nwakoleneh, O., Okwuonu, G., Clarnupsagoon, A., Pal, S., Parks, K.,
Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C.,
Plopper, F., Polinder, A., Popovic, D., Primus, E., Pu, L.-L.,
Puzo, M., Quirez, J., Rachlin, E., Reeves, K., Regier, M. A., Reigh, R.,
Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F.,
Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S. J.,
Sanders, W., Savery, G., Scherer, S., Scott, G., Shatsman, S., Shen, H.,
Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C. D., Smalls, D.,
Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Soza, J.,
Steinle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C.,
Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K.,
Valdes, R., Vera, V., Villalana, D., Waldron, L., Walker, B., Wang, J.,
Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F.,
Williams, G., Williams, R., Wlezyk, R., Wooden, H., Worley, K.,
Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,
Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von
Niederhausern, A., Weiser, R., Smith, D. R., Holt, R. A., Smith, H. O.,
Weinstock, G. and Gibbs, R. A.

TITLE
JOURNAL
REFERENCE
2 (bases 1 to 238379)
Unpublished
Worley, K. C.
TITLE
JOURNAL
Direct Submission
Submitted (12-JAN-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 238379)
Rat Genome Sequencing Consortium.
TITLE
JOURNAL
Direct Submission
Submitted (13-MAY-2003) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA

COMMENT
On May 13, 2003 this sequence version replaced gi:23269032.
The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described

```
Db 20757 GTCATTATGTGGATGACGATGGGAATGAGCTAGGCTCTGGTGTATATGGAGCTCAGAC 20816
QY 33 ThrGluLeuIleuTyrThrArgLysArgAspSerValLysTrpHisTyrLeuCysLeu 52
Db 20817 ACGAGCTGATTCTGTATACCCGGAACCGGACTCGGTGAAGTGGCACTACCTCTGCCTA 20876
QY 53 ArgArgTyrGlyTyrAspSerAsnLeuPheSerPheGluSerGlyArgCysGlnThr 72
Db 20877 CGACGATATGGCTATGACTCAAACTCTATTTCTTTGAAAGTGGTCGAAGGTGTCAAAC 20936
QY 73 GlyGlnGly----- 75
Db 20937 GGACAAGGTAGGTAAAGCCTTTATTTCTCCAGATGATACATATTCAAAGCACTGCACTTC 20996
QY 76 -----IlePheAlaPheLys-----CysAlaArgAlaGluGluLeu 87
Db 20997 CGGTACTCTTTGCATAAAGCCTGGCATGGAGACCCATGCTTGTACACCTAGCACTTG 21056
QY 88 PheAsnMetLeuGlnGluIleMetGlnAsnAsnSerIleAsnValValGluGluProVal 107
Db 21057 GGAGGTAGAGCAGGAGGATGTTGTAAATTTCTACATAGCAAATTCCTGGCAAGCCAGGG 21116
QY 108 ValGluArgAsnAsn-----HisGlnThr 115
Db 21117 TTAAGAAGGAGACACTATGCCAATAAATACCACACAACAACT 21158
```

Search completed: August 8, 2004, 23:07:18
Job time : 4938.38 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 8, 2004, 19:43:05 ; Search time 513.158 Seconds
(without alignments)
1076.210 Million cell updates/sec

Title: US-09-757-415a-1_COPY_11_140
Perfect score: 696
Sequence: 1 DIVPDNRHKKFKVINDDG.....NNHQTELEVPRTPTTPG 130

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delp 6.0, Delext 7.0

Searched: 337863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame_plus_p2n.model -DEV=xlh
-Q=/cgn2/USPRO.spool/US09757415/runat_02082004_101230_6675/app_query.fasta_1.526
-DB=N_Geneseq_29Jan04 -QFMT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPEXT=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09757415 -CGN 1.1.586 @runat_02082004_101230_6675 -NCPU=6 -ICPU=3
-NO_WMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Geneseq_29Jan04:
1: Geneseq1980s:
2: Geneseq1990s:
3: Geneseq2000s:
4: Geneseq2001as:
5: Geneseq2001bs:
6: Geneseq2002s:
7: Geneseq2003as:
8: Geneseq2003bs:
9: Geneseq2003cs:
10: Geneseq2004s:
11: Geneseq1980s:
12: Geneseq1990s:
13: Geneseq2000s:
14: Geneseq2001as:
15: Geneseq2001bs:
16: Geneseq2002s:
17: Geneseq2003as:
18: Geneseq2003bs:
19: Geneseq2003cs:
20: Geneseq2004s:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	696	100.0	1532	9	ADC35143 Human bre
2	696	100.0	2074	4	ABA09066 Human FGF
3	517.5	74.4	766	2	AAX39606 Breast ca
4	264.5	38.0	452	8	ACH40178 Human foe
5	174	25.0	1329	4	ABL22607 Drosophil
6	174	25.0	3329	4	ABL22606 Drosophil
7	143.5	20.6	1446	2	AAV38415 DNA encod
8	143.5	20.6	1972	9	ADD14695 Human src

9	143.5	20.6	3809	6	ABK95303 Human pro
10	142	20.4	205	2	AAV89621 EST clone
11	142	20.4	205	2	AAV89621 EST clone
12	141	20.3	1539	3	AAZ94091 Haematopo
13	141	20.3	1650	3	AAZ94125 Haematopo
14	137.5	19.8	1123	4	AAI60748 Human pol
15	137.5	19.8	1123	4	AAI60748 Human pol
16	137.5	19.8	1146	4	AAI58962 Human pol
17	137.5	19.8	1146	4	AAI58962 Human pol
18	137.5	19.8	1750	8	ACF08235 Human NOV
19	135.5	19.5	1221	3	AAAS2637 Eosinophi
20	134.5	19.3	2241	4	ABLO4039 Drosophil
21	134.5	19.3	5347	4	ABLO4038 Drosophil
22	133.5	19.2	914	4	AAI58963 Human pol
23	133.5	19.2	914	4	ADB48945 Novel hum
24	127.5	18.3	849	5	ABA12147 Human ner
25	127.5	18.3	1952	4	AAI59684 Human pol
26	124.5	17.9	1951	4	AAI57898 Human pol
27	124.5	17.9	2058	9	ADD71184 Human int
28	124.5	17.9	2253	4	AAH14202 Human cdn
29	124.5	17.9	2701	4	AAI60736 Human pol
30	124.5	17.9	2735	4	AAI58950 Human pol
31	124.5	17.9	2735	8	ADB48932 Novel hum
32	123.5	17.7	2805	4	AAH32551 Human sec
33	122	17.5	421	8	ACH47613 Human inf
34	116.5	16.7	403	5	AAF65334 Novel hum
35	113	16.2	508	8	ACH37206 Human end
36	111.5	16.0	1385	4	AAH34891 Human col
37	111.5	16.0	2181	4	AAH32617 Human sec
38	95.5	13.7	1409	6	ABQ54285 Human ova
39	95.5	13.7	30904	4	AAK75071 Human imm
40	92	13.2	373	4	AAK58197 Human imm
41	90.5	13.0	1743	3	AAAC9522 Human sec
42	90.5	13.0	1743	7	ABZ67189 Human sec
43	90.5	13.0	1743	7	ABZ73594 Secreted
44	90.5	13.0	1743	7	ADA98101 Human sec
45	90.5	13.0	1743	7	ADA43970 Human sec

ALIGNMENTS

RESULT 1
ADC35143
ID ADC35143 standard; cdna; 1532 BP.
XX AC ADC35143;
XX AC
XX AC
XX 18-DEC-2003 (first entry)
XX DT
XX DT
XX DE Human breast cancer antigen polynucleotide seq id 27.
XX XX breast cancer; breast cancer diagnosis; breast cancer antigen; gene; ss.
XX OS Homo sapiens.
XX PN US2003108888-A1.
XX XX
XX PD 12-JUN-2003.
XX XX
XX PF 15-MAY-2002; 2002US-00146473.
XX XX
XX PR 15-MAY-2001; 2001US-0291150P.
XX XX
XX PA (LUDW-) LUDWIG INST CANCER RES.
XX XX
XX PI Scanlan MJ, Gout I, Stockert E, Old LJ, Gure A, Chen Y;
XX XX WPI; 2003-829397/77.
XX DR P-PSDB; ADC35101.
XX XX
XX PT Diagnosing breast cancer in subject by obtaining biological sample from
XX PT subject, contacting sample with breast cancer-associated polypeptides,
XX PT determining specific binding between polypeptides and agents in sample.

Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-09-757-415A-1_COPY_11_140 (1-130) x ABA09066 (1-2074)

QY 1 AspThrValProAspAsnHisArgAsnLysPheLysValIleAsnValAspAspGly 20
DB 575 GACACTGTCCAGATAACCATCGAACAAGTTTAAGTCAATTAATGTGATGATGGG 634

QY 21 AsnGluLeuGlySerGlyIleMetGluLeuThrAspThrGluLeuIleLeuTyrThrArg 40
DB 635 AATGAGTGTAGTTCTGGCATATGGAACCTTACAGACAGAACTGATTTTATACACCCGC 594

QY 41 LysArgAspSerValIlystrPheHisTyrLeuCysLeuArgArgTyrGlyTyrAspSerAsn 60
DB 695 ARAAGTGACTCAGTAAATGGCACTACCTCTGCTCGACGCTATGGCTATGACTCGAAT 754

QY 61 LeuPheSerPheGluSerGlyArgArgCysGlnThrGlyGlnGlyIlePheAlaPheLys 80
DB 755 CTCCTTTTCTTTGAAAGTGTGGAAGGTGTCAAACTGGACAGGAATCTTTGCTTTTAAG 814

QY 81 CysAlaArgAlaGluGluLeuPheAsnMetLeuGlnGluIleMetGlnAsnAsnSerIle 100
DB 815 TGTGCCCGTGCAGAGAAATTTTAACTGTTTCAAGAGATTATGCAAAATAATAGTATA 874

QY 101 AsnValValGluGluProValValGluArgAsnHisGlnThrGluLeuGluValPro 120
DB 875 AATGTGTGTGAAGAGCGAGTTGTAGAAAGAAATAATCATCAGACAGAAATGGGAAGTCCCT 934

QY 121 ArgThrProArgThrProThrProGly 130
DB 935 AGAACACCTCGAACACCTACAACTCCAGGA 964

RESULT 3
AAAX39606
ID AAX39606 standard; DNA; 766 BP.

XX AC AAX39606;
XX DT 02-JUL-1999 (first entry)
XX DE Breast cancer associated gene.
XX KW Cancer associated antigen; diagnosis; research; treatment; human;
XX KW Breast cancer; colon cancer; gastric cancer; renal cancer; lung cancer;
XX KW prostate cancer; ss.
XX OS Homo sapiens.
XX PN W09904265-A2.
XX PD 28-JAN-1999.
XX PF 15-JUL-1998; 98WO-US014679.
XX PR 17-JUL-1997; 97US-00896164.
XX PR 10-OCT-1997; 97US-0061599P.
XX PR 10-OCT-1997; 97US-0061765P.
XX PR 10-OCT-1997; 97US-00948705.
XX PR 11-OCT-1997; 97GB-00021657.
XX PR 22-JUN-1998; 98US-00102322.
XX PA (LUDW-) LUDWIG INST CANCER RES.
XX PI Old LJ, Scanlan MJ, Stockert E, Gure A, Chen Y, Gout I;
XX PI O'hare M, Obata Y, Pfreundschuh M, Tureci O, Sahin U;
XX WPI; 1999-132448/11.
XX PT New isolated cancer associated nucleic acids and polypeptides - isolated
XX PT using sera from cancer patients, used to develop products for the
XX PT diagnosis, monitoring or treatment of cancers.

PS Claim 67; Page 375; 787pp; English.
XX The invention relates to a method for diagnosing a disorder characterised
CC by expression of a human cancer associated antigen precursor coded for by
CC a nucleic acid molecule (NAM). The method comprises: (a) contacting a
CC biological sample isolated from a subject with an agent that specifically
CC binds to the NAM, an expression product or a fragment of an expression
CC product complexed with an HLA molecule; and (b) determining the
CC interaction between the agent and the NAM or the expression product as a
CC determination of the disorder. The products and methods can be used in
CC the diagnosis, monitoring, research, or treatment of conditions
CC characterised by the expression of various cancer associated antigens.
CC The invention provides nucleic acid sequences and encoded polypeptides
CC which are cancer associated antigen precursors expressed in human breast
CC cancer, renal cancer, colon cancer, gastric cancer, prostate cancer and
CC lung cancer
XX Sequence 766 BP; 251 A; 151 C; 173 G; 190 T; 0 U; 1 Other;

Alignment Scores: Length: 766
Pred. No.: 5,33e-59 Matches: 120
Score: 517.50 Conservative: 2
Percent Similarity: 87.77% Mismatches: 7
Best Local Similarity: 86.33% Indels: 10
Query Match: 74.35% Gaps: 1
DB: 2

US-09-757-415A-1_COPY_11_140 (1-130) x AAX39606 (1-766)

QY 1 AspThrValProAspAsnHisArgAsnLysPheLysValIleAsnValAspAspGly 20
DB 359 GACACTGTCCAGATAACCATCGAACAAGTTTAAGTCAATTAATGTGATGATGGG 418

QY 21 AsnGluLeuGlySerGlyIleMetGluLeuThrAspThrGluLeuIleLeuTyrThrArg 40
DB 419 AATGAGTGTAGTTCTGGCATATGGAACCTTACAGACAGAACTGATTTTATACACCCGC 478

QY 41 LysArgAspSerValIlystrPheHisTyrLeuCysLeuArgArgTyrGlyTyrAspSerAsn 60
DB 479 ARAAGTGACTCAGTAAATGGCACTACCTCTGCTCGACGCTATGGCTATGACTCGAAT 538

QY 61 LeuPheSerPheGluSerGlyArgArgCysGlnThrGlyGlnGlyIlePheAlaPheLys 80
DB 539 CTCCTTTTCTTTGAAAGTGTGTCAAAGGTGTCAACTGGACAGGAATCTTTGCTTTTAA 598

QY 80 CysAlaArgAlaGluGluLeuPheAsnMetLeuGlnGluIleMetGlnAsnAsnSe 99
DB 599 GTGTGCCCGTGCAGAAAGAAATTTTAACTGTTTCAAGAGATTATGCAAAATAATAG 658

QY 99 rIleAsnValValGluGluProValValGluArgAsnHisGlnThrGluLeu 117
DB 659 TATAATGTGTGGAAGAGGCAATTTGTAGAAAGAAATAATCATCAGACAGAAATTTG 718

QY 118 GluValPro---ArgThrProArgThrProThrPro 129
DB 719 GAAGTCCCTAANNACAACTCGAACAACACTCAACTCCA 757

RESULT 4
ACH40178
ID ACH40178 standard; cDNA; 452 BP.
XX AC ACH40178;
XX DT 13-OCT-2003 (first entry)
XX DE Human foetal brain cDNA #1545.
XX KW Human; ss; sequencing by hybridisation; SBH; expressed sequence tag; EST;
XX KW genome mapping; biodiversity; genetic disorder.
XX OS Homo sapiens.
XX PN US2003073623-A1.

XX PD 17-APR-2003.
 XX PF 30-JUL-2001; 2001US-00918995.
 XX PR 30-JUL-2001; 2001US-00918995.
 XX PA (DRMA/) DRMANAC R T.
 XX PA (LABA/) LABAT I.
 XX PA (STAC/) STACHE-CRAIN B.
 XX PA (DICK/) DICKSON M C.
 XX PA (JONE/) JONES J W.
 XX PI Drmanac RT, Labat I, Stache-Crain B, Dickson MC, Jones LW;
 XX WPI; 2003-615964/58.
 XX New polynucleotide sequences obtained from various cDNA libraries, useful
 PT as hybridization probes, as oligomers for PCR, for chromosome and gene
 PT mapping, in the recombinant production of protein, or in generating
 PT antisense DNA or RNA.
 XX Claim 1; SEQ ID NO 27390; 44pp; English.
 XX The invention relates to an isolated polynucleotide comprising any one of
 CC 38043 cDNA sequences, appearing as ACH12789-ACH50831, whose sequence was
 CC determined by the technique of SSH (sequencing by hybridisation). Also
 CC included is a purified polypeptide comprising a sequence corresponding to
 CC a reading frame of the novel polynucleotide. The nucleic acid sequences
 CC are useful in diagnostics as expressed sequence tags (EST) for
 CC identifying expressed genes or for physical mapping of the human genome,
 CC in forensics, in assessing biodiversity, or in identifying mutations,
 CC responsible for genetic disorders and other traits. The nucleotide
 CC sequences are also useful as hybridisation probes, as oligomers for PCR,
 CC for chromosome and gene mapping, in the recombinant production of
 CC protein, or in generating antisense DNA or RNA. The purified polypeptide
 CC is useful for generating antibodies specific for it. The present sequence
 CC is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data
 CC for this patent did not form part of the printed specification, but was
 CC obtained in electronic format directly from USPTO at
 CC seqdata.uspto.gov/sequence.html?docID=20030073623
 XX SQ Sequence 452 BP; 98 A; 152 C; 108 G; 93 T; 0 U; 1 Other;

Alignment Scores:
 Pred. No.: 1,39e-25 Length: 452
 Score: 264.50 Matches: 49
 Percent Similarity: 84.72% Conservative: 12
 Best Local Similarity: 68.06% Mismatches: 10
 Query Match: 38.00% Indels: 1
 DB: .8 Gaps: 1

US-09-757-415A-1_COPY_11_140 (1-130) x ACH40178 (1-452)

QY 60 AsnLeuPheSerPheGluSerGlyArgArgCysGlnThrGlyGlnGlyIlePheAlaPhe 79
 Db 42 CACCTCTCTCTTGTAGAGTGGCCGCGATGTCAGACAGCCATGGAATATTGCAATT 101
 QY 80 LysCysAlaArgAlaGluGluLeuPheAsnMetLeuGlnGluIleMetGlnAsnAsnSer 99
 Db 102 AAGTGTTCGGGGCTGAGGAATCTTCAACCTCTTCAGGATCTGATGTCAGTGCACACG 161
 QY 100 IleAsnValValGluGluProVal---ValGluArgAsnAsnHisGlnThrGluLeuGlu 118
 Db 162 ATCAATGTGATGAAGAGCTGTTCATCATCCCGCAATAGCACCCCGCTGAGCTTGAC 221
 QY 119 ValProArgThrProArgThrProThrThrProGly 130
 Db 222 CTCCTCGAGCCCGCCAGCCACCAATGCTCTAGGC 257

RESULT 5
 ABL22607
 ID ABL22607 standard; DNA; 1329 BP.

XX AC ABL22607;
 XX DT 26-MAR-2002 (first entry)
 XX DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 19294.
 XX KW Drosophila; developmental biology; cell signalling; insecticide;
 XX KW pharmaceutical; gene; ds.
 XX OS Drosophila melanogaster.
 XX PN WO200171042-A2.
 XX PD 27-SEP-2001.
 XX 23-MAR-2001; 2001WO-US009231.
 XX PF 23-MAR-2000; 2000US-0191637P.
 XX PR 11-JUL-2000; 2000US-00614150.
 XX (PEKE) PE CORP NY.
 XX Venter JC, Adams M, Li PWD, Myers EW;
 XX WPI; 2001-656860/75.
 XX New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signaling and cell-cell
 PT interactions.
 XX Claim 1; SEQ ID NO 19294; 21pp + Sequence Listing; English.
 XX The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signaling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
 CC sequences (ABL01840-ABL16175) and the encoded proteins (ABBS7737-
 CC ABB2072). The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX SQ Sequence 1329 BP; 320 A; 385 C; 346 G; 278 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 7.51e-13 Length: 1329
 Score: 174.00 Matches: 42
 Percent Similarity: 52.46% Conservative: 22
 Best Local Similarity: 34.43% Mismatches: 52
 Query Match: 25.00% Indels: 6
 DB: 4 Gaps: 2

US-09-757-415A-1_COPY_11_140 (1-130) x ABL22607 (1-1329)

QY 7 HisArgAsnLysPheLysVal-----IleAsnValAspAspGlyAsnGlu 22
 Db 46 CATGAGAATGTTTCGAGTCCGAGTGCCTCACCTGCACCCGCCGGGTACGCCCATC 105
 QY 23 LeuGlySerGlyIleMetGluLeuThrAspThrGluLeuIleLeuTyThrArgLysArg 42
 Db 106 ATCCGAGTGTCTACCTGGAGCTAACCCCGGGAATAATATTCAGTCGCTGGGTGT 165
 QY 43 AspSerValLysTyrHisTyrLeuCysLeuArgArgTyrGlyTyrAspSerAsnLeuPhe 62
 Db 166 GAGCCCATGTCTGGGCACCTGCACACCTCCGACGTTATGGGTGTAATACGATCTCTC 225
 QY 63 SerPheGluSerGlyArgArgCysGlnThrGlyGlnGlyIlePheAlaPheLysCysAla 82
 Db 226 TCCTTCGAAGCAGGACGTAGTGTCTGTGGCCAGGATCTACACGTTTCGAGTCCAC 285
 QY 83 ArgAlaGluGluLeuPheAsnMetLeuGlnGluIleMetGlnAsnAsnSerIleAsnVal 102

```
Db 286 AATGCCGACGCTGTATCCGATGTTTCAGCGTACATCAACGCGGTGAATACAGATGCC 345
Qy 103 valGluGluProValValGluArgAsn-----HisGlnThrGluLeuGluValPro 120
Db 346 TTTGTGCGAGGGCAGCAGAGAAAGGTTAACTCGCTCATTCGGTGTCCGTAATAATATGGGC 405
Qy 121 ArgThr 122
Db 406 CGTACT 411

RESULT 6
ID ABL22606/c
XX ABL22606 standard; DNA; 3329 BP.
AC ABL22606;
XX
DT 26-MAR-2002 (first entry)
DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 19291.
XX
DE Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical; gene; ds.
XX
OS Drosophila melanogaster.
XX
XX WO200171042-A2.
XX
PD 27-SEP-2001.
XX
XX 23-MAR-2001; 2001WO-US009231.
XX
XX 23-MAR-2000; 2000US-0191637P.
XX
XX 11-JUL-2000; 2000US-00614150.
XX
XX (PEKE ) PE CORP NY.
XX
XX Venter JC, Adams M, Li FWD, Myers EW;
XX
XX WPI; 2001-658860/75.
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signaling and cell-cell
PT interactions.
XX
PS Claim 1; SEQ ID NO 19291; 2lpp + Sequence Listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (AB16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins (AB57737-
CC ABR72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 3329 BP; 880 A; 746 C; 747 G; 956 T; 0 U; 0 Other;
SQ

Alignment Scores:
Pred. No.: 2,76e-12 Length: 3329
Score: 174.00 Matches: 42
Percent Similarity: 52.46% Conservative: 22
Best Local Similarity: 34.43% Mismatches: 52
Query Match: 25.00% Indels: 6
DB: 4 Gaps: 2

US-09-757-415A-1_COPY_11_140 (1-130) x ABL22606 (1-3329)

Qy 7 HisArgAsnLysPheLysVal-----lleAsnValAspAspGlyAsnGlu 22
Db 2284 CATGAGATGTGTTTCGAGTGGAGTGGCTCACCTGTCAACCCACCCCGGTACGCCCATC 2225
```

```
Qy 23 LeuGlySerGlyIleMetGluLeuThrAspThrGluLeuLeuLeuLeuThrArgLysArg 42
Db 2224 ATCCGAGTGGCTACTCGAGCTAACGCCCGGGAACATAATATCCAGTCGCTGGTGT 2165
Qy 43 AspSerValLysTrpHisTyrLeuCysLeuArgArgTyrGlyTyrAspSerAsnLeuPhe 62
Db 2164 GAGCCCATTTGCTGGGCACCTCGACCTCGACGTTATGGTTGAATTAACCATCTCTTC 2105
Qy 63 SerPheGluSerGlyArgArgCysGlnThrGlyGlnGlyIlePheAlaPheLysCysAla 82
Db 2104 TCCCTCGAAGCAGGACGCTAGGTGTATGCTGGCCAGGATCTACACGTTTCGAGTCCAC 2045
Qy 83 ArgAlaGluGluLeuPheAsnMetLeuGlnGluLeuMetGlnAsnAsnSerIleAsnVal 102
Db 2044 AATGCCGAGCAGCTGTATCCGATGTTTCAGCGCTACATCAACGCGGTGAATACAGATGCC 1985
Qy 103 ValGluGluProValValGluArgAsnAsn-----HisGlnThrGluLeuGluValPro 120
Db 1984 TTTGTGAGGGGGAACGAGAAGGGTTAACTCCGCTCATTCGGTCTCGTAAATATGGGC 1925
Qy 121 ArgThr 122
Db 1924 CGTACT 1919

RESULT 7
AAV38415
ID AAV38415 standard; DNA; 1446 BP.
XX
AC AAV38415;
XX
DT 16-SEP-1998 (first entry)
DE DNA encoding p62 protein downstream of tyrosine kinases (p62-dok).
XX
KW p62-dok; protein downstream of tyrosine kinases;
KW tyrosine phosphorylation; receptor tyrosine kinase; oncogenic cell;
KW p120 ras GTPase-activating protein; GAP; signal transduction pathway;
KW antibody; diagnosis; treatment; chronic myelogenous leukaemia; CML;
KW human cancer; ss.
XX
OS Mammalia.
XX
XX Key Location/Qualifiers
PH 1. .1446
FT CDS /tag= a
FT /product= "p62-dok"
XX
XX WO9820129-A1.
XX
XX 14-MAY-1998.
XX
XX 30-OCT-1997; 97WO-US019788.
XX
XX 01-NOV-1996; 96US-0030418P.
XX
XX 22-JAN-1997; 97US-00787091.
XX
XX (COLD-) COLD SPRING HARBOR LAB.
XX (SLOK ) SLOAN KETTERING INST CANCER RES.
XX
XX Carpino NA, Kobayashi R, Wisniewski DG, Strife AO, Clarkson BD;
XX
XX WPI; 1999-286945/25.
XX
XX P-PSDB; AAW62507.
XX
XX Newly isolated human p62-dok protein - used to develop products for the
XX diagnosis and treatment of conditions in which oncogenic tyrosine kinase
XX is active, e.g. Chronic myelogenous leukaemia.
XX
XX Claim 1; Fig 2A-C; 86pp; English.
XX
XX The present sequence encodes p62-dok (protein downstream of tyrosine
XX kinases) protein. The protein is tyrosine phosphorylated in haematopoietic
CC
```

CC mammalian cells after stimulation with a ligand for a receptor tyrosine
 CC kinase found in oncogenic cells. The p62-dok protein binds p120 ras
 CC GTPase-activating protein (GAP). The p62-dok protein is involved in a
 CC signal transduction pathway initiated by receptor tyrosine kinase, and is
 CC rapidly tyrosine phosphorylated. The p62-dok protein is also present in
 CC normal cells, but is not constitutively phosphorylated by normal receptor
 CC tyrosine kinase. Antibodies against p62-dok and aberrantly phosphorylated
 CC p62-dok, and probes derived from the p62-dok DNA sequence can be used in
 CC the diagnosis and treatment of conditions in which an oncogenic tyrosine
 CC kinase is active such as chronic myelogenous leukaemia (CML) and other
 CC human cancers

XX SQ Sequence 1446 BP; 285 A; 457 C; 440 G; 264 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 9,88e-09 Length: 1446
 Score: 143.50 Matches: 33
 Percent Similarity: 47.32% Conservative: 20
 Best Local Similarity: 29.46% Mismatches: 50
 Query Match: 20.62% Indels: 9
 DB: 2 Gaps: 2

US-09-757-415A-1_COPY_11_140 (1-130) x AAV38415 (1-1446)

QY 24 GlySerGlyIleMetGluLeuThrAspThrGluLeuIleLeuTyrThrArgLysArgAsp 43
 Db 514 GGCTCCTACGTGCTGAGGTGGAGGCTGAAGGCTGACTCTCCTGACCGTGGGGGCCAG 573
 QY 44 Ser-----VallySTpHisTyrLeuCysLeuArgArgTyrGlyTyr 57
 Db 574 AGTCAGATACGAGGCCACTCTGCTTCGCGCCCTACACTCTGCTGCTGCTATGGCCGG 633
 QY 58 AspSerAsnLeuPheSerPheGluSerGlyArgArgCysGlnThrGlyGlnGlyIlePhe 77
 Db 634 GACAAAGGTCATGTTCTCTTCGAGCGCGCGCGCTGCCCTCAGCGCCCTGGAACTTC 693
 QY 78 AlaPheLysCysAlaArgAlaGluGluLeuPheAsnMetLeuGlnGluIleMetGlnAsn 97
 Db 694 ACCTTCCAGACGGCACAGGGAATGACATCTTCCAGGCGAGTGTGATCCATCCACCGG 753
 QY 98 AsnSerIleAsnValGlu-----GluProValValGluArgAsnAsnHisGln 114
 Db 754 CAGAAAGCCAGGGAAGAGCGGACAGCGGACACCATGTTCTCAGAGTGACTCCCATGAA 813
 QY 115 ThrGluLeuGluValProArgThrProArgThrPro 126
 Db 814 GGGAGGTGGCAGAGGGGAGTGTGCTTCCCACT 849

RESULT 8

ID ADD14695 standard; cDNA; 1972 BP.

XX AC ADD14695;

DT 01-JAN-2004 (first entry)

DE Human src biomarker polynucleotide SEQ ID NO:89.

XX predictor set; protein tyrosine kinase activity modulator;
 KW protein tyrosine kinase pathway; protein tyrosine kinase; cytosolic;
 KW gene therapy; drug sensitivity; genetic profile; cancer; human; gene; ss.

OS Homo sapiens.

XX W02003062395-A2.

XX 31-JUL-2003.

XX 17-JAN-2003; 2003WO-US001981.

XX 18-JAN-2002; 2002US-0350061P.

XX (BRIM) BRISTOL-MYERS SQUIBB CO.

XX Huang F, Fairchild CR, Lee FY, Shaw P;
 XX WPI; 2003-636735/60.

DR P-PSDB; ADD14099.

XX New polynucleotides and polypeptides for predicting the activity of
 PT compounds that interact with protein tyrosine kinases and/or protein
 PT tyrosine kinase pathways.

XX Claim 2; SEQ ID NO 89; 139pp; English.

XX The present invention describes a predictor set comprising a plurality of
 CC polynucleotides or polypeptides whose expression pattern is predictive of
 CC the response of cells to treatment with a compound that modulates protein
 CC tyrosine kinase activity or members of the protein tyrosine kinase
 CC pathway. Also described: (1) predicting whether a compound is capable of
 CC modulating the activity of cells, comprising obtaining a sample of cells,
 CC determining whether the cells express a plurality of markers, and
 CC correlating the expression of the markers to the compound's ability to
 CC modulate the activity of the cells; (2) a plurality of cell lines for
 CC identifying polynucleotides and polypeptides whose expression levels
 CC correlate with compound sensitivity or resistance of cells associated
 CC with a disease state; and (3) identifying polynucleotides and
 CC polypeptides that predict compound sensitivity or resistance of cells
 CC associated with a disease state, comprising subjecting the plurality of
 CC cell lines to one or more compounds, analysing the expression pattern of
 CC a microarray of polynucleotides or polypeptides, and selecting
 CC polynucleotides or polypeptides that predict the sensitivity or
 CC resistance of cells associated with a disease state by using the
 CC expression pattern of the microarray. The polynucleotides and
 CC polypeptides have cytostatic activities, and can be used in gene therapy.
 CC The polynucleotides and polypeptides are useful in predicting the
 CC activity of compounds that interact with protein tyrosine kinases and/or
 CC protein tyrosine kinase pathways. These may be used in determining drug
 CC sensitivity in patients to allow the development of individualized
 CC genetic profiles which aid in treating diseases and disorders (e.g.
 CC cancer) based on patient response at a molecular level. The present
 CC sequence is used in the exemplification of the present invention.

XX SQ Sequence 1972 BP; 415 A; 586 C; 623 G; 348 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 1.53e-08 Length: 1972
 Score: 143.50 Matches: 33
 Percent Similarity: 47.32% Conservative: 20
 Best Local Similarity: 29.46% Mismatches: 50
 Query Match: 20.62% Indels: 9
 DB: 2 Gaps: 2

US-09-757-415A-1_COPY_11_140 (1-130) x ADD14695 (1-1972)

QY 24 GlySerGlyIleMetGluLeuThrAspThrGluLeuIleLeuTyrThrArgLysArgAsp 43
 Db 600 GGCTCCTACGTGCTGAGGTGGAGGCTGAAGGCTGACTCTCCTGACCGTGGGGGCCAG 659

QY 44 Ser-----VallySTpHisTyrLeuCysLeuArgArgTyrGlyTyr 57
 Db 660 AGTCAGATACGAGGCCACTCTGCTTCGCGCCCTACACTCTGTTGCTGCTATGGCCGG 719

QY 58 AspSerAsnLeuPheSerPheGluSerGlyArgArgCysGlnThrGlyGlnGlyIlePhe 77
 Db 720 GACAAAGGTCATGTTCTCTTCGAGCGCGCGCGCTGCCCTCAGCGCCCTGGAACTTC 779

QY 78 AlaPheLysCysAlaArgAlaGluGluLeuPheAsnMetLeuGlnGluIleMetGlnAsn 97
 Db 780 ACCTTCCAGACGGCACAGGGAATGACATCTTCCAGGCGAGTGTGATCCATCCACCGG 839

QY 98 AsnSerIleAsnValGlu-----GluProValValGluArgAsnAsnHisGln 114

Db 840 CAGAAAGCCAGGGAAGAGCGGACAGGGGACCATGTTCTCAGAGTGACTCCCATGAA 899

QY 115 ThrGluLeuGluValProArgThrProArgThrPro 126

CC polynucleotide, which is a secreted EST, and the encoded protein are
 CC predicted to have useful biological activities which would make them
 CC suitable for treating, preventing or ameliorating medical conditions in
 CC humans and animals, although no supporting data is given. Suggested
 CC activities include nutritional activity, immune stimulating or
 CC suppressing activity, haematopoiesis regulating activity, tissue growth
 CC activity, activin/inhibin activity, chemotactic/chemokinetic activity,
 CC haemostatic and thrombolytic activity, receptor/ligand activity, anti-
 CC inflammatory activity, cadherin/tumour invasion suppressor activity,
 CC tumour inhibition activity. The polynucleotide may also be useful for
 CC gene therapy

XX Sequence 205 BP; 53 A; 30 C; 40 G; 82 T; 0 U; 0 Other;
 SQ

Alignment Scores:
 Pred. No.: 9.82e-10 Length: 205
 Score: 142.00 Matches: 28
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 20.40% Indels: 0
 DB: 2 Gaps: 0

US-09-757-415A-1_COPY_11_140 (1-130) x AAV89621 (1-205)

QY 75 GlyIlePheAlaPheLysCysAlaArgAlaGluGluLeuPheAsnMetLeuGlnGluile 94
 Db 113 GGAATCTTTGGCTTTAAAGTGTCGCCGTCGACAGAGAAATTTAACTGTTGCAAGAGATT 172
 QY 95 MetGlnAsnAsnSerIleAsnVal 102
 Db 173 ATGCAAAATAATAGTATAAATGTG 196

RESULT 11
 AAV89905
 ID AAV89905 standard; cDNA; 205 BP.

XX AC AAV89905;
 XX DT 15-FEB-1999 (first entry)
 XX DE EST clone CT474.
 XX KW Human; secreted protein; expressed sequence tag; EST; haematopoiesis;
 KW tissue growth; activin; inhibin; chemotaxis; chemokinesis; haemostatic;
 KW receptor; ligand; thrombolytic; anti-inflammatory; cadherin; anti-tumour;
 KW gene therapy; ss.
 XX OS Homo sapiens.
 XX PN WO9845436-A2.
 XX PD 15-OCT-1998.
 XX PF 10-APR-1998; 98WO-US006955.
 XX PR 10-APR-1997; 97US-00838821.
 XX PA (GEM) GENETICS INST INC.
 XX PI Jacobs K, McCoy JM, Lavallie ER, Racie LA, Merberg D, Treacy M;
 PI Spaulding V, Agostino MJ;
 XX WPI; 1999-070077/06.

XX New polynucleotides encoding human secreted proteins - derived from e.g.
 PT human blood, kidney, fetal lung, placenta, testes, brain, ovary,
 PT pituitary, retina and colon cDNA libraries.
 XX Claim 1; Page 368; 618pp; English.
 XX The present sequence represents a human expressed sequence tag (EST). The
 CC polynucleotide, which is a secreted EST, and the encoded protein are
 CC predicted to have useful biological activities which would make them

CC suitable for treating, preventing or ameliorating medical conditions in
 CC humans and animals, although no supporting data is given. Suggested
 CC activities include nutritional activity, immune stimulating or
 CC suppressing activity, haematopoiesis regulating activity, tissue growth
 CC activity, activin/inhibin activity, chemotactic/chemokinetic activity,
 CC haemostatic and thrombolytic activity, receptor/ligand activity, anti-
 CC inflammatory activity, cadherin/tumour invasion suppressor activity,
 CC tumour inhibition activity. The polynucleotide may also be useful for
 CC gene therapy

XX Sequence 205 BP; 53 A; 30 C; 40 G; 82 T; 0 U; 0 Other;
 SQ

Alignment Scores:
 Pred. No.: 9.82e-10 Length: 205
 Score: 142.00 Matches: 28
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 20.40% Indels: 0
 DB: 2 Gaps: 0

US-09-757-415A-1_COPY_11_140 (1-130) x AAV89905 (1-205)

QY 75 GlyIlePheAlaPheLysCysAlaArgAlaGluGluLeuPheAsnMetLeuGlnGluile 94
 Db 113 GGAATCTTTGGCTTTAAAGTGTCGCCGTCGACAGAGAAATTTAACTGTTGCAAGAGATT 172
 QY 95 MetGlnAsnAsnSerIleAsnVal 102
 Db 173 ATGCAAAATAATAGTATAAATGTG 196

RESULT 12
 AAZ94091
 ID AAZ94091 standard; cDNA; 1539 BP.

XX AC AAZ94091;
 XX DT 19-JUN-2000 (first entry)
 XX DE Haematopoietic stem cell specific nucleic acid SA61.
 XX KW SA61; haematopoietic stem cell; immune system disorder; leukaemia;
 KW antileukaemic; immunomodulator; therapy; mouse; ss.
 XX OS Mus musculus.
 XX PN WO200011168-A2.
 XX PD 02-MAR-2000.
 XX PF 20-AUG-1999; 99WO-US019052.
 XX PR 21-AUG-1998; 98US-00138132.
 XX PA (UYPR-) UNIV PRINCETON.
 XX PI Lemischka I, Moore K;
 XX WPI; 2000-237650/20.

XX Hematopoietic stem cell signaling proteins modulating replication and
 PT differentiation for treating immune system disorders and leukemia.
 XX Claim 10; Page 204; 256pp; English.
 XX The present sequence is that of nucleic acid SA61 isolated from mouse
 CC primitive stem cells by methods of the invention. It is an example of
 CC claimed isolated nucleic acids (see AAZ94077-294131) that are
 CC specifically expressed in haematopoietic stem cells (HSCs) and which
 CC encode HSC-specific proteins. The HSCs are especially primitive HSCs
 CC (PHSCs) such as umbilical cord cells, bone marrow cells and foetal liver
 CC cells. The encoded proteins (see AAY79176-93) are growth factors,
 CC transcription factors, splicing factors, capping factors, transport
 CC proteins, translation factors or replication factors that modulate HSC

CC activity, especially differentiation or replication. The invention
CC provides a claimed method for identifying PHSC-specific nucleic acids
CC involving: creating a PHSC cDNA library and a non-PHSC immune cell
CC library; and subtracting the 2 libraries. Also claimed are methods: for
CC generating a stem cell/progenitor cell from PHSCs; for identifying the
CC presence of a PHSC in a sample; for identifying the presence in a sample
CC of a compound that modulates HSC activity; for using such a compound to
CC treat an immune system condition, especially leukaemia; for introducing
CC exogenous nucleic acid into a HSC; and for ex vivo expansion of HSCs.
CC Also claimed is a PHSC specifically expressing 1 of the claimed nucleic
CC acids, such as the present sequence
XX
SQ Sequence 1539 BP; 319 A; 479 C; 463 G; 278 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 2,33e-08 Length: 1539
Score: 141.00 Matches: 36
Percent Similarity: 44.09% Conservative: 20
Best Local Similarity: 28.35% Mismatches: 47
Query Match: 20.26% Indels: 24
DB: 3 Gaps: 3

US-09-757-415A-1_COPY_11_140 (1-130) x AA294091 (1-1539)

QY 23 LeuGlySerGlyIleMetGluLeuThrAspThrGluLeuLeuLeuTyrThrArgLysArg 42
DB 560 CTGGCCCAAGATGACATCCCACTGAGGAGACA-----TCCAGCCCCAG 604
QY 43 AspSerValIysTrpHisTyrLeuCysLeuArgArgTyrGlyTyrAspSerAsnLeuPhe 62
DB 605 GCCTGTTTAACTGGCCCTACCGTTCTCTGCCCAAGTACGGCTCTGACCAAGGGTGTGTTTC 664
QY 63 SerPheGluSerGlyArgCysGlnThrGlyGlnGlyIlePheAlaPheLysCysAla 82
DB 665 TCGTTTGAAGCTGGCCGCGCTGTGACTCAGGTAGAGGCGCTTTTGCCTTCAGTACCCCG 724
QY 83 ArgAlaGluGluLeuPheAsnMetLeuGlnGluIleMetGlnAsnAsnSerIleAsnVal 102
DB 725 CFTGCCCCAGACATATGTGGGTTGTGGCTGCCGCCATTGCCCGCAGCGGAGCGTCTT 784
QY 103 ValGlu----- 104
DB 785 CCAGAGCTGGCCATGTCCCCACCGCTGCCCGCCCTCTCCCTGGCCCTCCCTCCCTA 844
QY 105 GluProValValGluArgAsnAsnHisGlnThrGluLeuGluValProArgThrProArg 124
DB 845 GAGCCCCCTGAGAGCTTCGGAGGTGGCCCCCAGGATTGAGCTGCCCATCCCAAGAG 904
QY 125 ThrPro---ThrThrProGly 130
DB 905 CTGCCTCTAAGTATCCCGGG 925

RESULT 13
AA294125
ID AA294125 standard; cDNA; 1650 BP.
XX AC AA294125;
XX DT 19-JUN-2000 (first entry)
XX DE Haematopoietic stem cell specific nucleic acid.
XX KW Haematopoietic stem cell; immune system disorder; leukaemia;
XX KW antileukaemic; immunomodulator; therapy; mouse; ss.
XX OS Mus musculus.
XX FH Location/Qualifiers
XX CDS 122..1204
XX FT /*tag= a
XX FT /*note= "in-frame stop codon at positions 1151-1153"
XX FN WO200011168-A2.

XX 02-MAR-2000.
XX 20-AUG-1999; 99WO-US019052.
XX 21-AUG-1998; 98US-00138132.
XX (UTPR-) UNIV PRINCETON.
XX Lemischka I, Moore K;
XX P-PSDB; AAY79184.
XX WPI; 2000-237650/20.
XX P-PSDB; AAY79184.
XX Haematopoietic stem cell signaling proteins modulating replication and
XX differentiation for treating immune system disorders and leukemia.
XX Claim 10; Page 231-232; 256pp; English.
XX The present sequence is that of a nucleic acid isolated from mouse
XX primitive stem cells by a method of the invention. It is an example of
XX claimed isolated nucleic acids (see AA294077-294131) that are
XX specifically expressed in haematopoietic stem cells (HSCs) and which
XX encode HSC-specific proteins. The HSCs are especially primitive HSCs
XX (PHSCs) such as umbilical cord cells, bone marrow cells and foetal liver
XX cells. The encoded proteins (see AAY79176-93) are growth factors,
XX transcription factors, splicing factors, capping factors, transport
XX proteins, translation factors or replication factors that modulate HSC
XX activity, especially differentiation or replication. The invention
XX provides a claimed method for identifying PHSC-specific nucleic acids
XX involving: creating a PHSC cDNA library and a non-PHSC immune cell
XX library; and subtracting the 2 libraries. Also claimed are methods: for
XX generating a stem cell/progenitor cell from PHSCs; for identifying the
XX presence of a PHSC in a sample; for identifying the presence in a sample
XX of a compound that modulates HSC activity; for using such a compound to
XX treat an immune system condition, especially leukaemia; for introducing
XX exogenous nucleic acid into a HSC; and for ex vivo expansion of HSCs.
XX Also claimed is a PHSC specifically expressing 1 of the claimed nucleic
XX acids, such as the present sequence
XX
SQ Sequence 1650 BP; 344 A; 511 C; 497 G; 298 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 2,57e-08 Length: 1650
Score: 141.00 Matches: 36
Percent Similarity: 44.09% Conservative: 20
Best Local Similarity: 28.35% Mismatches: 47
Query Match: 20.26% Indels: 24
DB: 3 Gaps: 3

US-09-757-415A-1_COPY_11_140 (1-130) x AA294125 (1-1650)

QY 23 LeuGlySerGlyIleMetGluLeuThrAspThrGluLeuLeuTyrThrArgLysArg 42
DB 671 CTGGCCCAAGATGACATCCCACTGAGGAGACA-----TCCAGCCCCAG 715
QY 43 AspSerValIysTrpHisTyrLeuCysLeuArgArgTyrGlyTyrAspSerAsnLeuPhe 62
DB 716 GCCTGTTTAACTGGCCCTACCGTTCTCTGCCCAAGTACGGCTCTGACCAAGGGTGTGTTTC 775
QY 63 SerPheGluSerGlyArgCysGlnThrGlyGlnGlyIlePheAlaPheLysCysAla 82
DB 776 TCGTTTGAAGCTGGCCGCGCTGTGACTCAGGTAGAGGCGCTTTTGCCTTCAGTACCCCG 835
QY 83 ArgAlaGluGluLeuPheAsnMetLeuGlnGluIleMetGlnAsnAsnSerIleAsnVal 102
DB 836 CFTGCCCCAGACATATGTGGGTTGTGGCTGCCGCCATTGCCCGCAGCGGAGCGTCTT 895
QY 103 ValGlu----- 104
DB 896 CCAGAGCTGGCCATGTCCCCACCGCTGCCCGCCCTCTCCCTGGCCCTCCCTA 955
QY 105 GluProValValGluArgAsnAsnHisGlnThrGluLeuGluValProArgThrProArg 124

Db 956 GAGCCCCCTGGAGAGCTTCGGGAGGTGGCCCCAGGATTTGAGTGGCTCCCAAG 1015
Qy 125 ThrPro---ThrThrProGly 130
Db 1016 CTGCCTCTAACTGATCCGGG 1036
RESULT 14
AAI60749
ID AAI60749 standard; cDNA; 1123 BP.
XX AC
XX AAI60749;
XX
XX 22-OCT-2001 (first entry)
XX
XX Human polynucleotide SEQ ID NO 4738.
XX
XX Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
XX peripheral nervous system; neuropathy; central nervous system; CNS;
XX Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
XX amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
XX chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
XX leukaemia; ss.
XX
XX Homo sapiens.
XX OS
XX WO200153312-A1.
XX PN
XX 26-JUL-2001.
XX PD
XX PF
XX 26-DEC-2000; 2000WO-US034263.
XX PR
XX 23-DEC-1999; 99US-00471275.
XX PR 21-JAN-2000; 2000US-00488725.
XX PR 25-APR-2000; 2000US-00552317.
XX PR 20-JUN-2000; 2000US-00598042.
XX PR 19-JUL-2000; 2000US-00620312.
XX PR 03-AUG-2000; 2000US-00653450.
XX PR 14-SEP-2000; 2000US-00662191.
XX PR 19-OCT-2000; 2000US-00693036.
XX PR 29-NOV-2000; 2000US-00727344.
XX
XX (HYSE-) HYSEQ INC.
XX PA
XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
XX Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;
XX Zhou P, Goodrich R, Drmanac RT;
XX
XX WPI: 2001-442253/47.
XX P-PSDB; AAM41593.
XX
XX Novel nucleic acids and polypeptides, useful for treating disorders such
XX as central nervous system injuries.
XX
XX Claim 1; SEQ ID NO 4738; 10078pp; English.
XX
XX The invention relates to human nucleic acids (AAI57798-AAI61369) and the
XX encoded polypeptides (AAM38642-AAM42213) with nootropic,
XX immunosuppressant and cytostatic activity. The polynucleotides are useful
XX in gene therapy. A composition containing a polypeptide or polynucleotide
XX of the invention may be used to treat diseases of the peripheral nervous
XX system, such as peripheral nervous injuries, peripheral neuropathy and
XX localised neuropathies and central nervous system diseases, such as
XX Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
XX lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
XX utilisation of the activities such as: immune system suppression,
XX Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
XX and thrombolytic activity, cancer diagnosis and therapy, drug screening,
XX assays for receptor activity, arthritis and inflammation, leukaemias and
XX C.N.S disorders. Note: The sequence data for this patent did not form
XX part of the printed specification
XX
XX Sequence 1123 BP; 337 A; 254 C; 283 G; 249 T; 0 U; 0 Other;

PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;
PI Zhou F, Goodrich R, Drmanac RT;
XX
DR WPI; 2001-442253/47.
DR P-PSDB; AAM41592.
XX
PT Novel nucleic acids and polypeptides, useful for treating disorders such
PT as central nervous system injuries.
XX
XX Claim 1; SEQ ID NO 4737; 10078pp; English.
XX
XX The invention relates to human nucleic acids (AAI57798-AAI61369) and the
CC encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC immunosuppressant and cytosstatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as; Immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders. Note: The sequence data for this patent did not form
CC part of the printed specification
XX
SQ Sequence 1123 BP; 337 A; 254 C; 283 G; 249 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 4.36e-08 Length: 1123
Score: 137.50 Matches: 34
Percent Similarity: 49.07% Conservative: 19
Best Local Similarity: 31.48% Mismatches: 42
Query Match: 19.76% Indels: 13
DB: 4 Gaps: 3

US-09-757-415A-1_COPY_11_140 (1-130) x AAI60748 (1-1123)
Qy 24 GlySerGlyIleMetGluLeuThrAspThrGluLeuIleLeuTyrThrArgLysArgAsp 43
Db 542 GGTGAATGCACATGCGATCATCATGAAATATCTATCTCTGGGATATCCACAATGCC 601
Qy 44 SerValIys-----TrrHisTyrLeuCysLeuArgArgTyrGlyTyrAspSerAsn 60
Db 602 AAGGTCAAACCTGGTGATGTGGCTCTCAGCTCACTGAGGACATACGGTCGGGACTCAACG 661
Qy 61 LeuPheSerPheGluSerGlyArgArgCysGlnThrClyGlnGlyIlePheAlaPheLys 80
Db 662 TGGTTCACGTTTGAGTCAGGAAGAAATGTGTGACACAGAGAGAGACTATTTCATTTTCAA 721
Qy 81 CysAlaArgAlaGluGluLeuPheAsnMetLeuGlnGluIleMetGlnAsnAsnSerIle 100
Db 722 ACAAGGGAAGGAGAA-----ATGATCTATCAGAAGGTTTCATTCTGGGACACTG 769
Qy 101 AsnValValGluGlu-----ProValValGluArgAsnAsnHisGln 114
Db 770 GCCATAGCTGAGCAACATGAAGATTAAATGCTAGAAATGGAACAGAGGCCCGGCTTCAG 829
Qy 115 ThrGluLeuGluValProArgThr 122
Db 830 ACAAGCTTGACTGACCAATGACA 853

Search completed: August 8, 2004, 21:31:31
Job time : 524.158 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

CM protein - nucleic search, using frame_plus_p2n model

Run on: August 8, 2004, 21:03:30 ; Search time 3606.64 Seconds
(without alignments)
1076.370 Million cell updates/sec

Title: US-09-757-415A-1_COPY_11_140
Perfect score: 696
Sequence: 1 DTVPDNRHNRKFKVINDDGG.....NNHQTELEVPRTPTPTTPG 130

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DBV=xlh
-Q=/cgn2/USPTO_spool/US09757415/runat_02082004_101231_6697/app_query.fasta_1.526
-DB=EST -QPMI=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR_SCORE=pcr -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09757415 @CGN 1.1 4385 @runat_02082004_101231_6697 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estm:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_man:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	685	98.4	2918	11	AK028813	CA751075 UI-M-FOO-
2	679	97.6	914	14	CA751075	CA751075 UI-M-FOO-
3	672	96.6	1527	29	AV411600	AV411600 Homo sapi
4	652	93.7	1517	29	AV411602	AV411602 Mus muscu
5	640	92.0	800	14	CF747997	CF747997 UI-M-HEO-
6	622.5	89.4	567	9	AL638785	AL638785 AL638785
7	622.5	89.4	721	9	AL647878	AL647878 AL647878
8	616.5	88.6	555	10	BF615429	BF615429 DG22D08.Y
9	616.5	88.6	689	13	EX848607	EX848607 BX848607
10	616.5	88.6	779	14	CF290491	CF290491 AGENCOURT
11	616.5	88.6	850	14	CA793389	CA793389 AGENCOURT
12	616.5	88.6	885	13	EQ735693	EQ735693 AGENCOURT
13	610.5	88.1	733	13	EU482662	EU482662 603368412
14	610.5	87.7	669	12	EU073369	EU073369 B0073369
15	609.5	87.6	639	9	AL866682	AL866682 AL866682
16	606.5	87.1	622	12	BJ621383	BJ621383 BJ621383
17	602.5	86.6	429	10	BE681241	BE681241 g080c07.Y
18	592	85.2	1527	29	AY411601	AY411601 pan tscgl
19	582.5	83.7	630	9	AL848080	AL848080 AL848080
20	559.5	80.4	567	12	BJ037146	BJ037146 BJ037146
21	546	78.4	636	9	AL859119	AL859119 AL859119
22	544	78.2	587	14	CB585347	CB585347 AMGNNUC:U
23	544	78.2	604	14	CB581691	CB581691 AMGNNUC:U
24	523	75.1	706	13	BU704334	BU704334 UI-M-FOO-
25	515	74.0	688	12	BU794794	BU794794 UMN03809
26	511.5	73.5	886	13	EX717562	EX717562 EX717562
27	510.5	73.3	917	14	CD557174	CD557174 AGENCOURT
28	507.5	72.9	776	14	CD109594	CD109594 AGENCOURT
29	501.5	72.1	448	14	CD730859	CD730859 4039849 1
30	489.5	70.3	857	14	CB983469	CB983469 AGENCOURT
31	488.5	70.2	783	12	BI757794	BI757794 603029751
32	485.5	69.8	837	14	CB960314	CB960314 AGENCOURT
33	484.5	69.6	753	13	EQ179061	EQ179061 UI-M-EVO-
34	482.5	69.3	720	12	EM951508	EM951508 UI-M-EGO-
35	482.5	69.3	797	10	BF338354	BF338354 602035593
36	478.5	68.7	629	12	BM741175	BM741175 K-EST0013
37	478	68.7	651	9	AL845765	AL845765 AL845765
38	474.5	68.2	723	10	AW957161	AW957161 EST369231
39	474.5	68.2	1201	13	EX331765	EX331765 BX331765
40	467.5	67.2	779	12	E1520005	E1520005 603071305
41	462	66.4	550	14	CB045562	CB045562 NISC_gc11
42	460.5	66.2	772	9	A1815668	A1815668 au49d04.Y
43	457	65.7	474	13	EX098956	EX098956 BX098956
44	455	65.4	542	10	BE112403	BE112403 UI-R-BJ1-
45	450	64.7	591	12	EG297442	EG297442 602395590

ALIGNMENTS

RESULT 1
AK028813
LOCUS
DEFINITION
2918 bp mRNA linear HTC 18-SEP-2003
Mus musculus 10 days neonate skin cDNA, RIKEN full-length enriched library, clone:4732458E18 product:SUCL1-ASSOCIATED NEUROTHIOPIC FACTOR TARGET (FGFR SIGNALLING ADAPTOR) homolog [Homo sapiens], full insert sequence.
ACCESSION
AK028813
VERSION
AK028813.1 GI:26324755
KEYWORDS
HTC; CAP trapper.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE

AUTHORS

Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)

JOURNAL

MEDLINE

PUBMED

PUBMED

PUBMED

PUBMED

PUBMED

PUBMED

PUBMED

PUBMED

PUBMED

PUBMED

PUBMED

PUBMED

PUBMED

PUBMED

PUBMED

PUBMED

PUBMED

PUBMED

PUBMED

PUBMED

PUBMED

PUBMED

PUBMED

PUBMED

PUBMED

PUBMED

PUBMED

PUBMED

PUBMED

PUBMED

PUBMED

PUBMED

PUBMED

PUBMED

PUBMED

PUBMED

PUBMED

PUBMED

PUBMED

PUBMED

PUBMED

PUBMED

PUBMED

PUBMED

PUBMED

PUBMED

PUBMED

PUBMED

PUBMED

PUBMED

PUBMED

PUBMED

PUBMED

PUBMED

PUBMED

PUBMED

PUBMED

PUBMED

PUBMED

PUBMED

PUBMED

PUBMED

PUBMED

PUBMED

PUBMED

PUBMED

PUBMED

PUBMED

PUBMED

PUBMED

PUBMED

PUBMED

PUBMED

PUBMED

PUBMED

PUBMED

PUBMED

PUBMED

PUBMED

PUBMED

PUBMED

PUBMED

PUBMED

PUBMED

PUBMED

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
10349636

2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Komno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20499374

3 Shibata, K., Itoh, M., Aizawa, K., Nagao, K., Sasaki, N., Carninci, P.,
Komno, H., Akiyama, J., Nishi, K., Kishimoto, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohata, E., Watahiki, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsumura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
Riken integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multichannel sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913

4 The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)
11076861

5 The PANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 2918)

7 Fukuda, S., Furuno, M., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,
Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,
Hori, F., Imocani, J., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,
Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N.,
Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,
Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,
Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,
Takeda, Y., Tanaka, T., Tomaru, A., Toyota, T., Yasunishi, A.,
Muramatsu, M. and Hayashizaki, Y.
Direct Submission
Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan [E-mail: genome-res@gsc.riken.go.jp],
URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222,
Fax: 81-45-503-9216

8 Encyclopaedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in Riken.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL: http://genome.gsc.riken.go.jp/
URL: http://fantom.gsc.riken.go.jp/
1. 2918

9 /organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"

10 /db_xref="PANTOM_DB:4732458E18"
/db_xref="MGI:2350817"
/db_xref="taxon:10090"
/clone="4732458E18"
/tissue_type="skin"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="10 days neonate"
372_-1898

11 /note="unnamed protein product: SUC1-ASSOCIATED
NEUTROTROPIC FACTOR TARGET (FGFR SIGNALING ADAPTOR)
homolog [Homo sapiens] (SPTF[AAH21562], evidence: FASTV,
95.5%id, 100%length, match=1536)
putative"
/codon_start=1
/protein_id="BAC26132.1"
/db_xref="GI:26324756"

12 /translation="MGSCCSPDKDTPVDNHRNFKVINUDDGNEIGSGWELTDTE
LILYTRKRSVWVHYLCURRYGIDSNLFESGRRRCQCGIFAFKARAELEENMLQ
EIMQNSINVEEPVRSVSHQTELEVPRTPTTGLGAQNPNGPYRPSFGDAS
SHPSRHPVSGARLPSVEESTPLVAEEQVHTYVNTTQVQERKRASVHVPDA
RVSAESTPSKEPSPEDRPQVLLKPEGVFLVPTVQKMERKELEQLGKDPV
SGSAGNTEWDTGVDSERRDVPVKNLVENINGLSIPASGVRRRLTSTSDTQ
NINNSAOPRELLAYENLPSLPVWEARKLSRDEDDNLGPKTSLNGVHNLDPMNY
VNTENVTPASAHKIDYSKRDCCTPTVFNDIRPSLEHRLNLYQVDSGSGSDNP
QTPKTPPTPLPPTPRTRTELYAVIDIETAAWSNLQKALPDDGTSRKTRHNSDLEPM
"

13 Alignment Scores:
Pred. No.: 1,63e-76 Length: 2918
Score: 685.00 Matches: 127
Percent Similarity: 100.00% Conservativity: 3
Best Local Similarity: 97.69% Mismatches: 0
Query Match: 98.42% Indels: 0
DB: 11 Gaps: 0

14 US-09-757-415A-1_COPY_11_140 (1-130) x AK028813 (1-2918)

15 QY 1 AspThrValProAspAsnHisArgAsnLysPhelysValIleAsnValAspAspGly 20
|||||
16 Db 402 GACACTGTCCAGATAACCATCGGAACAAAGTTTAAGTTCATTAAATGATGATGATG 461
|||||

17 QY 21 AsnGluLeuGlySerGlyIleMetGluLeuThrAspThrGluLeuLeuLeuThrArg 40
|||||
18 Db 462 AATGAGTAGGCTCTGGTGATGAGTCACTACACACACAGAGCTGATCTGTACACCGG 521
|||||

19 QY 41 LysArgAspSerValLysTyrHisTyrLeuCysLeuArgArgTyrGlyTyrAspSerAsn 60
|||||

20 Db 522 AAACGGGACTCGGTAATAATGACACTACCTCTGCTAGCAGATACGGCTATGACTCAAT 581
|||||

21 QY 61 LeuPheSerPheGluSerGlyArgArgCysGlnThrGlyGlnGlyIlePheAlaPheLys 80
|||||

22 Db 582 CCGTTCCTTTTGAAGGTCGGAAGGTGTCAGACTGGACAAAGAAATTTTTCCTTTTAA 641
|||||

23 QY 81 CysAlaArgAlaGluGluLeuPheAsnMetLeuGlnGluMetGlnAsnAsnSerIle 100
|||||

24 Db 642 TGTGCCCGTGCGAAGAAATTTTAAACATGTTGCAAGAGATTATGCAAAATAATAGTATA 701
|||||

25 QY 101 AsnValValGluGluProValValGluArgAsnAsnHisGlnThrGluLeuValPro 120
|||||

26 Db 702 AATGGTGGGAAGACCGGTTGTGAAAGAGGATGATCATCAGACAGAGCTGGAAGTCCCT 761
|||||

27 QY 121 ArgThrProArgThrProThrProGly 130
|||||

28 Db 762 AGGACACCTCGGACACCTACACTCCAGG 791
|||||

RESULT 2

CA751075

LOCUS

DEFINITION

CA751075

ACCESSION

VERSION

CA751075.1

GI:25577987

914 bp

mRNA

linear

EST 09-JUL-2003

UI-M-FOO-cdm-p-10-0-UI.r1 NIH BMAP_F00 Mus musculus cDNA clone

IMAGE:6830555 5', mRNA sequence.

KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 914)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapsb@mail.nih.gov
Tissue Procurement: Dr. Jim Lin, University of Iowa
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)

Seq primer: pYX-5.
FEATURES
source Location/Qualifiers
1..914
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:6830555"
/tissue_type="whole brain"
/dev_stage="embryo 12.5dpc"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH BMAP_F00"
/notes="Organ: Brain; Vector: pYX-Asc; Site 1: EcoR I;
Site 2: Not I; The library was constructed according
Bonaldo, Lennon and Soares, Genome Research, 6:791-806.
1996. Denatured RNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with oligo-dT
primer containing a Not I site. Double strand cDNA was
size selected according to mRNA size fraction, ligated
with EcoR I adaptor, digested with NotI and then cloned
directly into pYX-Asc vector. The library tag
sequence located between the Not I site and the polyA tail
is "TGAGAGGCC. This library was created for the University
Iowa Brain Anatomy Project (BMAP): 'Gene discovery in the
Developing Mouse Nervous System', supported by National
Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,
program coordinator."

ORIGIN

Alignment Scores:
Pred. No.: 1,68e-76 Length: 914
Score: 679.00 Matches: 126
Percent Similarity: 99.23% Conservative: 3
Best Local Similarity: 96.92% Mismatches: 1
Query Match: 97.56% Indels: 0
DB: 14 Gaps: 0

US-09-757-415A-1_COPY_11_140 (1-130) x CA751075 (1-914)

QY 1 AspThrValProAspAsnHisArgAsnLysPheLysValIleAsnValAspAspGly 20
DB 423 GACACTGTCCAGATACCAATCGGAACAGTTTAAAGTCATTAAATGCGATGATGGG 482
QY 21 AsnGluLeuGlySerGlyIleMetGluLeuThrAspThrGluLeuIleuTyThrArg 40
DB 483 AATGAGTAGCGCTCGTGAATGGCTACCTCTGCTACGACGATACGCTGATCAAT 542
QY 41 LysArgAspSerValLysTrpHisLsTyLeuCysLeuArgArgTyGlyTyRaspSerAsn 60
DB 543 AAACGGGACTCGTAAATGGCTACCTCTGCTACGACGATACGCTGATCAAT 602
QY 61 LeuPheSerPheGluSerGlyArgArgCysGluThrGlyGlnGlyIlePheAlaPheLys 80

DB 603 CTGTTTCTTTTGAAGCTGGTGCAGAGGTGTCAGACTCGACAAGGAATTTTTCCTTTTAAAG 662
QY 81 CysAlaArgAlaGluGluLeuPheAsnMetLeuGlnGluIleMetGlnAsnAsnSerIle 100
DB 663 TGTCCCGTGCAGAGAAGTAATTATTAAACATGTTGCAAGAGATTATGCANAAATAATAGTATA 722
QY 101 AsnValValGluGluProValValGluArgAsnAsnHisGlnThrGluLeuGluValPro 120
DB 723 AATGTGGTGGAGAGCCGGTTGTAGAAAGAGTAGTATCATCAGACAGAGCTGGAGTCCCT 782
QY 121 ArgThrProArgThrProThrThrProGly 130
DB 783 AGGACACCTCGACACCTTACAACTCCAGGG 812
RESULT 3
LOCUS AY411600 1527 bp DNA linear GSS 12-DEC-2003
DEFINITION Homo sapiens HCM4278 gene, VIRTUAL TRANSCRIPT, partial sequence,
genomic survey sequence.
ACCESSION AY411600
VERSION AY411600.1 GI:39767568
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1527)
AUTHORS Clark,A.G., Glancowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferrera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 1527)
AUTHORS Clark,A.G., Glancowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferrera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
source Location/Qualifiers
1..1527
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
<1..>1527
/locus_tag="HCM4278"
Alignment Scores:
Pred. No.: 2.9e-75 Length: 1527
Score: 672.00 Matches: 126
Percent Similarity: 96.92% Conservative: 0
Best Local Similarity: 96.92% Mismatches: 4
Query Match: 96.55% Indels: 0
DB: 29 Gaps: 0
US-09-757-415A-1_COPY_11_140 (1-130) x AY411600 (1-1527)
QY 1 AspThrValProAspAsnHisArgAsnLysPheLysValIleAsnValAspAspGly 20
DB 31 GACACTGTCCAGATACCAATCGGAACAGTTTAAAGTCATTAAATGCGATGATGGG 90
QY 21 AsnGluLeuGlySerGlyIleMetGluLeuThrAspThrGluLeuIleuTyThrArg 40
DB 91 AATGAGTTAGGTTCTGGCATATGGAATTACAGACAGACTGATTTTATACACCGC 150
QY 41 LysArgAspSerValLysTrpHisTyLeuCysLeuArgArgTyGlyTyRaspSerAsn 60

```

151 AAAGTACTCAGTAAATGGCACTACTCTGCTCGAGCTATGGCTATGACTGAAT 210
61 LeuPheSerPheGluSerGlyArgArGysGlnThrGlyGlnGlyLeuPheAlaPheLys 80
211 CTCCTTTCTTTTGAAGTGGTGAAGGTGCAAACTGGACAAGGAATCTTTGCTTTAAG 270
81 CysAlaArgAlaGluGluLeuPheAsnMetLeuGlnGluLeuMetGlnAsnSerIle 100
271 TGTCCCGCTGCAGAAGAAATTAATTAATGATGTCAGAGATTAATGCAAAATAATAGTATA 330
101 AsnValValGluGluProValValGluArgAsnHisGlnThrGluLeuGluValPro 120
331 AATGTGGTGAAGACCCAGTGTAGAAAGAAATAATCATCAGACAGAAATTTGGAAGTCCT 390
121 ArgThrProArgThrProThrThrProGly 130
391 AGAACACCTCGANNNNNNNNNNCTCCAGGA 420

RESULT 4
LOCUS AY411602 1517 bp DNA linear GSS 12-DEC-2003
DEFINITION Mus musculus HCM4278 gene, VIRTUAL TRANSCRIPT, partial sequence,
genomic survey sequence.
ACCESSION AY411602
VERSION AY411602.1 GI:39767570
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 1517)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarilwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 1517)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarilwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
source
Location/Qualifiers
1..1517
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
<1..>1517
/locus_tag="HCM4278"

ORIGIN
Alignment Scores:
Pred. No.: 1.08e-72 Length: 1517
Score: 652.00 Matches: 124
Percent Similarity: 90.07% Conservative: 3
Best Local Similarity: 87.94% Mismatches: 2
Query Match: 93.68% Indels: 12
DB: 29 Gaps: 1

US-09-757-415a-1_COPY_11_140 (1-130) x AY411602 (1-1517)

Qy 1 AspThrValProAsnHisArgAsnLysPheLysValIleAsnValAspAspGly 20
31 GACACTGCCAGTATACCATCGGAACAGTTTAAGTCAATTAATGATGATGATGATG 90

```

```

Qy 21 AsnGluLeuGlySerGlyIleMetGluLeuThrAspThrGluLeuIleLeuTyThrArg 40
Db 91 AATGAGCTAGGCTCTGGTGATGATGAACCTACAGACACAGAGCTGATTCGTACACCGG 150
Qy 41 LysArgAspSerValIleTTPHisTyrLeuCysLeuArgArgTyrGlyTyrAspSerAsn 60
Db 151 AAACGGGACTCGGTAATAATGCACTACCTCTGCTCAGCAGATACGGCTATGACTCAAT 210
Qy 61 LeuPheSerPheGluSerGlyArgArGysGlnThrGlyGlnGlyLeuPheAlaPheLys 80
Db 211 CTGCTTTCTTTTGAAGTGGTGAAGGTGTCAGACTGGACAAGGAATTTTTCCTTTAAG 270
Qy 81 CysAlaArgAlaGluGluLeuPheAsnMetLeuGlnGluLeuMetGlnAsnSerIle 100
Db 271 TGTCCCGCTGCAGAAGAAATTAATTAATGATGTCAGAGATTAATGCAAAATAATAGTATA 330
Qy 101 AsnValValGluGluProValValGluArgAsnHisGlnThrGluLeuGluValPro 120
Db 331 AATGTGGTGAAGACCCAGTGTAGAAAGGAAGTAGTATCATCAGACAGAGCTGGAAGTCCT 390
Qy 121 ArgThrProArg-----ThrProThrThr 128
Db 391 AGGACACCTTCGGTCCAGGGCTCGGTCTCAGAACTTACCTAATGGGTATCCCGCTACC 450
Qy 129 Pro 129
Db 451 CCT 453

RESULT 5
LOCUS CF747997 800 bp mRNA linear EST 10-OCT-2003
DEFINITION UI-M-HE0-cme-d-20-0-UI.r1 NIH BMAP HE0 Mus musculus cDNA clone
IMAGE:30629755 5', mRNA sequence.
ACCESSION CF747997
VERSION CF747997.1 GI:37644341
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 800)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Dr. James Lin Unive
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at
http://genome.uiowa.edu/distribution/mouseefi.html
This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)
Seq primer: pyx-5.
Location/Qualifiers
1..800
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:30629755"
/tissue_type="whole eye"
/dev_stage="embryo 12.5, 13.5, 14.5 dpc"
/lab_hosts="DH10B (T1 phage resistant)"
/clone_lib="NIH_BMAP_HE0"
/note="Organ: Eye; Vector: pyx- Asc; Site 1: EcoR I;
Site 2: Not I; The library was constructed according
Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
1996. Denatured RNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with oligo-dT
primer containing a Not I site. Double strand cDNA was
size selected according to mRNA size fraction, ligated

```


with EcoR I adaptor , digested with NotI and then cloned directionally into pYX-Asc vector . The library tag sequence located between the Not I site and the polyA tail is TTTATCAAGT. This library was created for the University of Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH)."

ORIGIN

Alignment Scores:
 Pred. No.: 1,45e-71 Length: 800
 Score: 640.00 Matches: 123
 Percent Similarity: 97.67% Conservatives: 3
 Best Local Similarity: 95.35% Mismatches: 3
 Query Match: 91.95% Indels: 1
 DB: 14 Gaps: 0

US-09-757-415A-1_COPY_11_140 (1-130) x CF747997 (1-800)

QY 1 AspThrValProAspAsnHisArgAsnLysPheLysValIleAsnValAspAspGly 20
 DB 415 GACACTGTCCAGATAACCATCGGAACAAGTTTAAGTCAATTAATGTGGATGATGGG 474
 QY 21 AsnGluLeuGlySerGlyIleMetGluLeuThrAspThrGluLeuIleLeuThrArg 40
 DB 475 ATGACTAGGCTCTGGTGTGATGGACTCAGACAGAGCTGATCTGTACACCGG 534
 QY 41 LysArgAspSerValLysTrpHisTyrLeuCysLeuArgArgTyrGlyTyrAspSerAsn 60
 DB 535 AAACGGGACTCGGTAAATAGGCACTACCTCTGCTACGACGATACGGCTATGACTCAAT 594
 QY 61 LeuPheSerPheGluSerGlyArgArgCysGlnThrGlyGlnGlyIlePheAlaPheLys 80
 DB 595 CTGTTTCTTTTGAAGTGGTGAAGGTGTGACACTGGACAGGAATTTTCTTTTAA 654
 QY 81 CysAlaArgAlaGluLeuPheAsnMetLeuGlnGlnIleMetGlnAsnAsnSerIle 100
 DB 655 TGTGCCCGGTGCAGACAATA-TTTAACAATGTTGCAAGAGATTATGCANAAATATAGTATA 713
 QY 101 AsnValValGluGluProValValGluArgAsnAsnHisGlnThrGluLeuValPro 120
 DB 714 AATGTCGTGGAAGCGGTGTGAGAAAGGAGTAGTATCATCAGACAGAGCTGGNAGTCCCT 773
 QY 121 ArgThrProArgThrProThrThrPro 129
 DB 774 AGCACCTCGGACACTCAACTCCA 800

RESULT 6

AL638785 567 bp mRNA linear EST 07-NOV-2001
 LOCUS AL638785 XGC-egg Silurana tropicalis cDNA clone L1EBb02 5', mRNA
 DEFINITION

ACCESSION AL638785.1 GI:16790764
 VERSION AL638785
 SOURCE Silurana tropicalis (western clawed frog)

ORGANISM Silurana tropicalis

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae; Xenopodinae; Silurana.
 1 (bases 1 to 567)

REFERENCE Huckle, E., Taylor, R., Ashurst, J.L., Zorn, A.M. and Rogers, J.
 Sanger Xenopus tropicalis EST project 2001 (10_2001)
 TITLE Unpublished (2001)
 JOURNAL

COMMENT Sequencing primer: SP6
 This sequence is from a Xenopus Gene Collection (XGC) library constructed by Aaron M. Zorn.

FEATURES

source
 1..567
 /organism="Silurana tropicalis"
 /mol_type="mRNA"
 /db_xref="taxon:8364"
 /clone="L1EBb02"
 /dev_stage="egg"
 /lab_host="Escherichia coli XL1-blue"
 /clone_lib="XGC-egg"
 /note="Vector: pCS107; Site 1: EcoRI; Site 2: NotI; cDNA was oligo dt primed from Sug of poly A+ RNA from egg. EcoRI-NotI cut cDNA was then ligated into pCS107 with EcoRI at the 5' end and NotI at the 3' end"

ORIGIN

Alignment Scores:
 Pred. No.: 1.55e-69 Length: 567
 Score: 622.50 Matches: 118
 Percent Similarity: 96.15% Conservatives: 7
 Best Local Similarity: 90.77% Mismatches: 4
 Query Match: 89.44% Indels: 1
 DB: 9 Gaps: 1

US-09-757-415A-1_COPY_11_140 (1-130) x AL638785 (1-567)

QY 1 AspThrValProAspAsnHisArgAsnLysPheLysValIleAsnValAspAspGly 20
 DB 103 GAATCTATCCAGATAACCAACAAACAGATTTAAGTTTATTAACGTTGATGATGGT 162
 QY 21 AsnGluLeuGlySerGlyIleMetGluLeuThrAspThrGluLeuIleLeuThrArg 40
 DB 163 AATGAGCTTGCTCTGGCATTAATGGAATTGCAGAAATAGCTAATCTTATACACTGT 222
 QY 41 LysArgAspSerValLysTrpHisTyrLeuCysLeuArgArgTyrGlyTyrAspSerAsn 60
 DB 223 AAGCGAGATTCTGTCAAGTGGCTTACCTTCGCTGGACCTATGATGATGATCAAC 282
 QY 61 LeuPheSerPheGluSerGlyArgCysGlnThrGlyGlnGlyIlePheAlaPheLys 80
 DB 283 CTATTTCTTTTGAAGTGGACGAGGTGTGACACTGGACAGGAATATTGCTTTTAA 342
 QY 81 CysAlaArgAlaGluGluLeuPheAsnMetLeuGlnGlnIleMetGlnAsnAsnSerIle 100
 DB 343 TGTGTCGTGCAGAGAGCTCTTCAATATGCTTCAAGAGATCATGCAGAAATAGCAT 402
 QY 101 AsnValValGluGluProValValGluArgAsnAsnHisGlnThrGluLeuValPro 120
 DB 403 AGCGTTGTGAGAGAGCCAGTAGTTGAACGA---AATCCCCAGACTGAGTTGAGTTCCA 459
 QY 121 ArgThrProArgThrProThrThrProGly 130
 DB 460 AGAACCCCTCGCACACCACCACCCTTGA 489

RESULT 7

AL647878

LOCUS

DEFINITION

AL647878 XGC-gastrula Silurana tropicalis cDNA clone TGAS039C07 5', mRNA sequence.

ACCESSION AL647878.2 GI:38458802

VERSION AL647878

KEYWORDS EST.

SOURCE Silurana tropicalis (western clawed frog)

ORGANISM Silurana tropicalis

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae; Xenopodinae; Silurana.

1 (bases 1 to 721)

REFERENCE Croning, M.D.R., Ashurst, J.L., Taylor, R., Zorn, A.M. and Rogers, J.

AUTHORS Sanger Xenopus tropicalis EST project 2001 (11_2003)

TITLE Unpublished (2003)

JOURNAL On Dec 13, 2001 this sequence version replaced gi:17656321.

COMMENT Contact: Huckle E

Sanger Institute

Hinxton, Cambridgeshire, CB10 1SA, UK

Email: tropesanger.ac.uk

This sequence is from a Xenopus Gene Collection (XGC) library constructed by Aaron M. Zorn.

cDNA was oligo dT primed from Sug of poly A+ RNA from stages 10-13 gastrulae. EcoRI-NotI cut cDNA was then ligated into pCS107 with EcoRI at the 5' end and NotI at the 3' end.

Vector: pCS107; Site 1: EcoRI; Site 2: NotI

Host: Escherichia coli XL1-blue

Sanger Xenopus tropicalis EST project 2001

TROPICALIS SEQUENCE ID: TGAS039C07.plkSP6

Sequencing primer: SP6

FEATURES

source

Location/Qualifiers

1..721

/organism="Silurana tropicalis"

/mol_type="mRNA"

/db_xref="taxon:8364"

/clone="TGAS039C07"

/dev_stage="gastrula (stages 10.5-12 mixed)"

/lab_host="Escherichia coli XL1-blue"

/clone_lib="XGC-gastrula"

/note=vector: pCS107; Site 1: EcoRI; Site 2: NotI; cDNA

was oligo dT primed from Sug of poly A+ RNA from stages

10-13 gastrulae. EcoRI-NotI cut cDNA was then ligated

into pCS107 with EcoRI at the 5' end and NotI at the 3'

end."

ORIGIN

Alignment Scores:

Pred. No.: 2,23e-69 Length: 721
Score: 622.50 Matches: 118
Percent Similarity: 96.15% Conservative: 7
Best Local Similarity: 90.77% Mismatches: 4
Query Match: 89.44% Indels: 1
DB: 9 Gaps: 1

US-09-757-415A-1_COPY_11_140 (1-130) x AL647878 (1-721)

QY 1 AspThrValProAspAsnHisArgAsnLysPheLysValIleAsnValAspAspGly 20
DB 260 GAACATATCCAGATACCAACAAACAGATTATAGTTTATTAACCTTGATGATGTT 319
QY 21 AsnGluLeuGlySerGlyIleMetGluLeuThrAspThrGluLeuLeuLeuThrArg 40
DB 320 AATGAGCTTGGCTCTGGCATATGAATGACAGAAATGAGCTAATCTATACACTCT 379
QY 41 LysArgAspSerValIysTrpHisTyrLeuCysLeuArgArgTyrGlyTyrAspSerAsn 60
DB 380 AAGCAGATTCTGTCAAGTGGCTTACCTCTGCTTGGAGCTATGATGATGATCAAC 439

QY 61 LeuPheSerPheGluSerGlyArgArgCysGlnThrGlyGlnGlyIlePheAlaPheLys 80
DB 440 CTATTTCTTTTGAAAGTGCACGAGGTGTACAGCTGGACAGGAATATTGCTTTTAA 499
QY 81 CysAlaArgAlaGluLeuPheAsnMetLeuGlnGluLeuMetGlnAsnSerIle 100
DB 500 TGTGCTGCTGCAGAGAGCTCTTCAATATGCTTCAAGAGATCATGCAGAACATGACATT 559
QY 101 AsnValValGluGluProValValGluArgAsnHisGlnThrGluLeuGluValPro 120
DB 560 AGCGTTGTGGAGGCGCAGTAGTTGAACGA--AATCCCCAGACTGAGTTGGATGTCCA 616

QY 121 ArgThrProArgThrProThrThrProGly 130
DB 617 AGAACCCCTCGCACACCCACCCCTCGGA 646

RESULT 8

BF615429

LOCUS

DEFINITION

GS22008.y1 Xenopus laevis gastrula non normalized Xenopus laevis

cDNA clone IMAGE:3749439 5' similar to TR:043558 O43558 FGFR

SIGNALING ADAPTOR SNT-1.1, mRNA sequence.

ACCESSION

BF615429

VERSION

KEYWORDS

SOURCE

ORGANISM

Xenopus laevis

Xenopus laevis

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;

Xenopodinae; Xenopus.

1 (bases 1 to 555)

Clifton, S., Johnson, S.L., Blumberg, B., Song, J., Hillier, L.,

Pape, D., Martin, J., Wylie, T., Underwood, K., Theising, B., Bowers, Y.,

Paxon, B., Gibbons, M., Harvey, N., Ritter, E., Jackson, Y., McCann, R.,

Waterston, R. and Wilson, R.

WashU Xenopus EST project, 1999

Unpublished (1999)

Contact: Sandy Clifton, Ph.D.

WashU Xenopus EST project, 1999

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

Library constructed by Bruce Blumberg

Library normalized by Jihwan Song

DNA Sequencing by: Washington University Genome Sequencing Center

Source lab clone id - xlnuga006h16 This clone is available

royalty-free through LNL; contact the IMAGE Consortium

(image@image.llnl.gov) for further information.

Seq primer: -40Rp from Gibco

High quality sequence stop: 419.

Location/Qualifiers

1..555

/organism="Xenopus laevis"

/mol_type="mRNA"

/db_xref="taxon:8355"

/clone="IMAGE:3749439"

/tissue_type="gastrula (stages 10.5, 11.5 mixed)"

/lab_host="Top-10 F"

/clone_lib="Xenopus laevis gastrula non normalized"

/note=vector: pBluescript SK-; Site 1: EcoRI; Site 2:

XhoI; cDNA was prepared from 2 ug of poly A+ RNA (equal

parts from stage 10.5 and stage 11.5 gastrulae).

EcoRI-XhoI cut cDNA was then ligated into UniLap-XR

(Stratagene) with EcoRI at the 5' end and XhoI at the 3'

end. The library was mass excised and used to infect

Top10F. Clones were picked into freezing medium (per

liter 15 g tryptone, 10g yeast extract, 5g NaCl, 36 mM

K2HPO4, 13.2 mM KH2PO4, 1.7 mM Na-citrate, 0.4 mM MgSO4, 7

H2O, 6.8 mM (NH4)2SO4, 4 % w/v glycerol) and grown for 24

hours. Original library construction by Bruce Blumberg

(Cho et al 1991 Cell 67, 1111-1120)."

ORIGIN

Alignment Scores:

Pred. No.: 8,89e-69 Length: 555
Score: 616.50 Matches: 116
Percent Similarity: 96.15% Conservative: 9
Best Local Similarity: 89.23% Mismatches: 4
Query Match: 88.58% Indels: 1
DB: 10 Gaps: 1

US-09-757-415A-1_COPY_11_140 (1-130) x BF615429 (1-555)

QY 1 AspThrValProAspAsnHisArgAsnLysPheLysValIleAsnValAspAspGly 20
DB 46 GAAACTATCCAGATACCAACAAACAGATTATAGCTTATTAATGTGTGATGATGTT 105
QY 21 AsnGluLeuGlySerGlyIleMetGluLeuThrAspThrGluLeuLeuLeuThrArg 40
DB 106 AATGAGCTTGGCTCTGGCATATGAATGACAGAAATGATCTCATCTTATACACTCG 165
QY 41 LysArgAspSerValIysTrpHisTyrLeuCysLeuArgArgTyrGlyTyrAspSerAsn 60
DB 166 AAGCAGATTCTGTCAAGTGGCTTACCTCTGCTTGGAGCTATGATGATGATCAAC 225

```

QY 61 LeuPheSerPheGluSerGlyArgAsnCysGlnThrGlyGlnGlyPheAlaPheLys 80
DB 226 CTTTTCCTTTTGAAGAGTGGACGAAGTGTGCAGCTGGACAGGATATTTGCTTTTAA 285
QY 81 CysAlaArgAlaGluLeuPheAsnMetLeuGlnGluLeuMetGlnAsnSerIle 100
DB 286 TGTGCCCGCGCAGAGAGCTCTTCAATATGCTTCAAGAGATCATGCAAAATATAGCAT 345
QY 101 AsnValValGluGluProValValGluArgAsnAsnHisGlnThrGluLeuValPro 120
DB 346 AGCTTGTGAGAGGCGAGTAGTTGACGA---AATCCCGAGACTGAGTTGATTTCCA 402
QY 121 ArgThrProArgThrProThrThrProGly 130
DB 403 AGAACCCCTCGCACCAACCAACCCCTGGA 432

RESULT 9
BX848607 689 bp mRNA linear EST 11-DEC-2003
LOCUS BX848607 NICHDXGC Emb4 Xenopus laevis cDNA clone IMAGE999M0812312
DEFINITION IMAGE:5570479 5', mRNA sequence.
ACCESSION BX848607
VERSION BX848607.1 GI:39737355
KEYWORDS EST.
SOURCE Xenopus laevis (African clawed frog)
ORGANISM Xenopus laevis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
Xenopodinae; Xenopus.
1 (bases 1 to 689)
REFERENCE Heil,O., Neubert,P., Peters,M., Radelof,U., Schneider,D.,
AUTHORS Schroth,A., Korn,B. and Landgrabe,J.
TITLE Xenopus laevis UniGene Set 1 (RZPDLIB No.988)
JOURNAL Unpublished (2003)
COMMENT RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
RZPDLIB; I.M.A.G.E. cDNA Clone Collection (amp- resistant) (RZPDLIB
No.998) http://www.rzpd.de/cgi-
bin/products/showlib.pl.cgi?response?libNo=998 RZPDLIB; Xenopus
laevis UniGene Set 1 (RZPDLIB No.988) http://www.rzpd.de/cgi-
bin/products/showlib.pl.cgi?response?libNo=988 Contact: Ina Rolfs
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Heubnerweg 6, D-14059 Berlin, Germany
Tel: +49 30 32639 101
Fax: +49 30 32639 111
www.rzpd.de
This clone is available royalty-free from RZPD;
contact RZPD (clone@rzpd.de) for further information. Seq primer:
SP6, 5' ATTAGTGACACTTAG 3'.

FEATURES
source
location/Qualifiers
1..689
/organism="Xenopus laevis"
/mol_type="mRNA"
/db_xref="taxon:8355"
/clone="IMAGE999M0812312 ; IMAGE:5570479"
/dev_stage="embryo, stage 31-32"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NICHDXGC Emb4"
/notes="Organ: whole embryo; Vector: pCMV-SPORT6; Site:1:
NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo
dT. Average insert size 2.1 kb. Constructed by Life
Technologies. Note: This is a Xenopus Gene Collection
(XGC) library."

ORIGIN
Alignment Scores:
Pred. No.: 1.23e-68 Length: 689
Score: 616.50 Matches: 117
Percent Similarity: 95.38% Conservative: 7
Best Local Similarity: 90.00% Mismatches: 5

```

```

Query Match: 89.58% Indels: 1
DB: 13 Gaps: 1
US-09-757-415a-1_copy_11_140 (1-130) x BX848607 (1-689)
QY 1 AspThrValProAspAsnHisArgAsnLysPheLysValIleAsnValAspAspGly 20
DB 204 GAAACTATCCAGATAACCAACAACAGAGATTTAAGGTTATTAACTGATGATGGT 263
QY 21 AsnGluLeuGlySerGlyIleMetGluLeuThrGluLeuLeuLeuThrArg 40
DB 264 AATGAGCTTGGCTCTGGCATATATGGAATTGAGAGAAATGACTCATCTTATACACTGC 323
QY 41 LysArgAspSerValLysTrpHisTyrIleuCysLeuArgArgTyrGlyTyrAspSerAsn 60
DB 324 AAGAGAGATTCTGTCAAGTGGCTTACCTCTGCTGGACGCTATGATAGATTCAAC 383
QY 61 LeuPheSerPheGluSerGlyArgCysGlnThrGlyGlnGlyIlePheAlaPheLys 80
DB 384 CTTTTTCTTTTGAAGAGTGGACGAAGTGTGCAGCTGGACAGGATATTTGCTTTTAA 443
QY 81 CysAlaArgAlaGluLeuPheAsnMetLeuGlnGluLeuMetGlnAsnSerIle 100
DB 444 TGTGCCCGCGCAGAGAGCTCTTCAATATGCTTCAAGAGATCATGCAAAATATAGCAT 503
QY 101 AsnValValGluGluProValValGluArgAsnAsnHisGlnThrGluLeuValPro 120
DB 504 AGTGTGTAGAGGACCCCTGTAGTTGAAACGA---AATCCCGAGCCGAGTTGGATTTCCA 560
QY 121 ArgThrProArgThrProThrThrProGly 130
DB 561 AGAACCCCTCGCACCAACCAACCCCTGCG 590

RESULT 10
CF290491 779 bp mRNA linear EST 14-AUG-2003
LOCUS CF290491 NICHDXGC Emb4 Xenopus laevis cDNA clone
DEFINITION IMAGE:5515552 5', mRNA sequence.
ACCESSION CF290491
VERSION CF290491.1 GI:33650270
KEYWORDS EST.
SOURCE Xenopus laevis (African clawed frog)
ORGANISM Xenopus laevis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
Xenopodinae; Xenopus.
1 (bases 1 to 779)
REFERENCE NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
TITLE Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics / NIH
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgapbs-remail.nih.gov
Tissue Procurement: Dr. Igor Dawid
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium
www-bio.llnl.gov/bbrp/image/image.html
Plate: LLAM12174 row: 1 column: 17
High quality sequence stop: 665.
location/Qualifiers
1..779
/organism="Xenopus laevis"
/mol_type="mRNA"
/db_xref="taxon:8355"
/clone="IMAGE:5515552"
/dev_stage="embryo, stage 31-32"
/lab_host="DH10B (phage-resistant)"

```

/clone_lib="NICHD_XGC_Emb4"
 /note="Organ: whole embryo; Vector: pCMV-SPORT6; Site 1:
 NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo
 dt. Average insert size 2.1 kb. Constructed by Life
 Technologies. Note: This is a Xenopus Gene Collection
 (XGC) library."

ORIGIN

Alignment Scores:
 Pred. No.: 1.48e-68 Length: 779
 Score: 616.50 Matches: 117
 Percent Similarity: 95.38% Conservative: 7
 Best Local Similarity: 90.00% Mismatches: 5
 Query Match: 88.58% Indels: 1
 DB: 14 Gaps: 1

US-09-757-415A-1_COPY_11_140 (1-130) x CF290491 (1-779)

QY 1 AspThrValProAspAsnHisArgAsnLysPheLysValIleAsnValAspAspGly 20
 Db 198 GAAACTATCCAGATACCAACAAACAGATTTAAGGTTATTACGTTGAATGATGGT 257
 QY 21 AsnGluLeuGlySerGlyIleMetGluLeuThrAspThrGluLeuIleLeuThrArg 40
 Db 258 AATGAGCTTGCTCTGCGCATATGGAATTGAGAGAAAATGAATCACTTATACACTGC 317
 QY 41 LysArgAspSerValLysTrpHisTyrLeuCysLeuArgArgTyrGlyTyrAspSerAsn 60
 Db 318 AAGAGAGATTCTGTCAAGTGGCTTACTCTGCTTGGAGCGTATGATACGATTCAAC 377
 QY 61 LeuPheSerPheGluSerGlyArgCysGlnThrGlyGlnGlyIlePheAlaPheLys 80
 Db 378 CTTTTCCTTTTGAAGTGGACGAGAGTGTGCAGCTGCACAGGAATATTGCTTTTAAG 437
 QY 81 CysAlaArgAlaGluGluLeuPheAsnMetLeuGlnGluIleMetGlnAsnSerIle 100
 Db 438 TGTGCCCGCGCAGAGAGCTCTTCAATATGCTTCAAGAGATCATGCAAAATATAGCAT 497
 QY 101 AsnValValGluGluProValValGluArgAsnHisGlnThrGluLeuGluValPro 120
 Db 498 AGTGTGTAGAGAGAGCTGTAGTTGAACGA---AATCCCGAGACCGAGTTGGATGTCCA 554
 QY 121 ArgThrProArgThrProThrThrProGly 130
 Db 555 AGAACCCCTCGCACACCCACTACCCCTGGC 584

RESULT 11

CA793389 850 bp mRNA linear EST 04-DEC-2002
 DEFINITION AGENCOURT 10256250 NICHD_XGC_Emb4 Xenopus laevis cDNA clone
 IMAGE:5157160 5', mRNA sequence.
 CA793389
 VERSION CA793389.1 GI:26042166
 KEYWORDS EST.
 SOURCE Xenopus laevis (African clawed frog)

ORGANISM
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
 Xenopodinae; Xenopus.

REFERENCE
 1 (bases 1 to 850)
 NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 Unpublished (1997)
 CONTACT: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Martha Rebbert, Steven L. Klein, Ph.D.
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>

Plate: LLM11389 row: 9 column: 17
 High quality sequence stop: 600.
 Location/Qualifiers
 1..850
 source

/organism="Xenopus laevis"
 /mol_type="mRNA"
 /db_xref="taxon:8355"
 /clone="IMAGE:5157160"
 /tissue_type="embryo (stage 10)"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NICHD_XGC_Emb4"
 /note="Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI;
 Cloned unidirectionally. Primer: Oligo dt. Average insert
 size 1.55 kb. Constructed by Life Technologies. Note: This
 is a Xenopus Gene Collection (XGC) library."

ORIGIN

Alignment Scores:
 Pred. No.: 1.69e-68 Length: 850
 Score: 616.50 Matches: 117
 Percent Similarity: 95.38% Conservative: 7
 Best Local Similarity: 90.00% Mismatches: 5
 Query Match: 88.58% Indels: 1
 DB: 14 Gaps: 1

US-09-757-415A-1_COPY_11_140 (1-130) x CA793389 (1-850)

QY 1 AspThrValProAspAsnHisArgAsnLysPheLysValIleAsnValAspAspGly 20
 Db 215 GAAACTATCCAGATACCAACAAACAGATTTAAGGTTATTACGTTGAATGATGGT 274
 QY 21 AsnGluLeuGlySerGlyIleMetGluLeuThrAspThrGluLeuIleLeuThrArg 40
 Db 275 AATGAGCTTGCTCTGCGCATATGGAATTGAGAGAAAATGAATCACTTATACACTGC 334
 QY 41 LysArgAspSerValLysTrpHisTyrLeuCysLeuArgArgTyrGlyTyrAspSerAsn 60
 Db 335 AAGAGAGATTCTGTCAAGTGGCTTACTCTGCTTGGAGCGTATGATACGATTCAAC 394
 QY 61 LeuPheSerPheGluSerGlyArgCysGlnThrGlyGlnGlyIlePheAlaPheLys 80
 Db 395 CTTTTCCTTTTGAAGTGGACGAGAGTGTGCAGCTGCACAGGAATATTGCTTTTAAG 454
 QY 81 CysAlaArgAlaGluGluLeuPheAsnMetLeuGlnGluIleMetGlnAsnSerIle 100
 Db 455 TGTGCCCGCGCAGAGAGCTCTTCAATATGCTTCAAGAGATCATGCAAAATATAGCAT 514
 QY 101 AsnValValGluGluProValValGluArgAsnHisGlnThrGluLeuGluValPro 120
 Db 515 AGTGTGTAGAGAGAGCTGTAGTTGAACGA---AATCCCGAGACCGAGTTGGATGTCCA 571
 QY 121 ArgThrProArgThrProThrThrProGly 130
 Db 572 AGAACCCCTCGCACACCCACTACCCCTGGG 601

RESULT 12

BQ735693
 LOCUS
 DEFINITION
 AGENCOURT 8149157 NICHD_XGC_Emb4 Xenopus laevis cDNA clone
 IMAGE:5570479 5', mRNA sequence.
 BQ735693
 VERSION BQ735693.1 GI:21874590
 KEYWORDS EST.

SOURCE
 ORGANISM
 Xenopus laevis (African clawed frog)

Xenopus laevis
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
 Xenopodinae; Xenopus.

REFERENCE
 1 (bases 1 to 885)
 NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 Unpublished (1997)

COMMENT

Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Dr. Igor David
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 www.bio.llnl.gov/bbrp/image/image.html
 Plate: LAM12312 row: m column: 08
 High quality sequence stop: 550.

FEATURES

source

1..885
 /organism="Xenopus laevis"
 /mol_type="mRNA"
 /db_xref="taxon:8355"
 /clone="IMAGE:5570479"
 /dev_stage="embryo, stage 31-32"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NICHD XGC Emb4"
 /notes="Organ: whole embryo; Vector: pCMV-SPORT6; Site 1:
 NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo
 dr. Average insert size 2.1 kb. Constructed by Life
 Technologies. Note: This is a Xenopus Gene Collection
 (XGC) library."

ORIGIN

Alignment Scores:

Pred. No.: 1..8e-68 Length: 885
 Score: 616.50 Matches: 117
 Percent Similarity: 95.38% Conservative: 7
 Best Local Similarity: 90.00% Mismatches: 5
 Query Match: 88.58% Indels: 1
 DB: 13 Gaps: 1

US-09-757-415a-1_copy_11_140 (1-130) x BQ735693 (1-885)

QY 1 AsnThrValProAspAsnHisArgAsnLysPheLysValIleAsnValAspAspGly 20
 Db 202 GAACATCTCCAGATACCAACAAACAGATTTAAGTTATTAACTGATGATGATG 261
 QY 21 AsnGluLeuGlySerGlyIleMetGluLeuThrAspThrGluLeuIleLeuTyrThrArg 40
 Db 262 AATGAGCTGGCTGGCATAATGGAATTCGAGAGAAATGAATCATCTTATACATCGC 321
 QY 41 LysArgAspSerValLysTrpHisTyrLeuCysLeuArgArgTyrGlyTyrAspSerAsn 60
 Db 322 AAGAGAGATTCTGTAAGTGGCTTACCTCTGCTTGGAGCGCTATGGATCAAC 381
 QY 61 LeuPheSerPheGluSerGlyArgArgCysGlnThrGlyGlnGlyIlePheAlaPheLys 80
 Db 382 CTCTTTTCTTTTGAAGTGACCAAGGTGTGAGACTCGACACAGGAATATTGCTTTTAA 441
 QY 81 CysAlaArgAlaGluGluLeuPheAsnMetLeuGlnGluIleMetGlnAsnSerIle 100
 Db 442 TGTGCCCGCGAGAGAGCTCTTCAATATGCTTCAAGAGATCATGCAAAATATACATT 501
 QY 101 AsnValValGluGluProValValGluArgAsnAsnHisGlnThrGluLeuGluValPro 120
 Db 502 AGTGTGTAGAGGAGCTGTAGTTGACGA--AATCCCCAGACCGAGTTGGATGTTCCA 558
 QY 121 ArgThrProArgThrProThrThrProGly 130
 Db 559 AGAACCCCTCGACACCCACTACCCCTGGC 588

RESULT 13

BU462662
 LOCUS 733 bp mRNA linear EST 29-NOV-2002
 DEFINITION 603368412f1 CSEOREN19 Gallus gallus cDNA clone CHEST27194 5', mRNA
 sequence.
 ACCESSION BU462662
 VERSION BU462662.1 GI:25951982
 KEYWORDS EST.

SOURCE

ORGANISM

Gallus gallus (chicken)
 Gallus gallus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
 Phasianinae; Gallus.

REFERENCE

AUTHORS

Boardman, P.E., Sans-Ezquerro, J., Overton, I.M., Burt, D.W., Bosch, E.,
 Pong, W.T., Tickle, C., Brown, W.R.A., Wilson, S.A. and Hubbard, S.J.
 A Comprehensive Collection of Chicken cDNAs
 Curr. Biol. 12 (22), 1965-1969 (2002)

TITLE

JOURNAL

MEDLINE

PUBMED

COMMENT

Contact: Simon Hubbard
 Department of Biomolecular Sciences
 University of Manchester Institute of Science and Technology
 (UMIST)
 PO Box 88, Manchester, M60 1QD, UK
 Tel: 01612008930
 Fax: 01612360409
 Email: Simon.Hubbard@umist.ac.uk.

FEATURES

source

1..733
 /organism="Gallus gallus"
 /mol_type="mRNA"
 /strain="Layer"
 /db_xref="taxon:9031"
 /clone="CHEST27194"
 /sex="Female"
 /dev_stage="adult"
 /lab_host="DH10B"
 /clone_lib="CSEOREN19"
 /note="Organ: ovary; Vector: pBluescript II KS(+); Site 1:
 EcoRI; Site 2: NotI; This normalized library was
 constructed from 1 million independent clones. cDNA
 synthesis was initiated using an oligo(dT) primer, using
 methylated C in the first strand reaction, double-stranded cDNA
 Following this first strand reaction, digested with
 EcoRI, size-selected, and cloned into the NotI and EcoRI
 compatible sites of a custom modified MCS of the
 pBluescript (KS+) vector. The library was normalized in 2
 rounds using conditions adapted from Soares et al., PNAS
 (1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6
 (1996): 791, except that a significantly longer
 reannealing hybridization was used."

ORIGIN

Alignment Scores:
 Pred. No.: 3..82e-68 Length: 733
 Score: 613.00 Matches: 119
 Percent Similarity: 98.36% Conservative: 1
 Best Local Similarity: 97.54% Mismatches: 2
 Query Match: 88.07% Indels: 1
 DB: 13 Gaps: 0

US-09-757-415a-1_copy_11_140 (1-130) x BU462662 (1-733)

QY 9 AsnLysPheLysValIleAsnValAspAspGlyAsnGluLeuGlySerGlyIleMet 28
 Db 4 CACAAGTTTAAAGTTATTAAATGATGATGATGATGATGATGATGATGATGATGATG 63
 QY 29 GluLeuThrAspThrGluLeuIleLeuTyrThrArgLysArgAspSerValLysTrpHis 48
 Db 64 GAACCTACAGATACCGAACTAAT-TTGTACACCCGTAAGAGGACTCTGTAAATGGCAC 122
 QY 49 TyrLeuCysLeuArgArgTyrGlyTyrAspSerAsnLeuPheSerPheGluSerGlyArg 68
 Db 123 TACCTCTGCTCCGTCGCTATGGCTACGACTCGAATCTTTCTCTTTTGAAGTGGTAGA 182
 QY 69 ArgCysGlnThrGlyGlnGlyIlePheAlaPheLysCysAlaArgAlaGluGluLeuPhe 88
 Db 183 AGGTGTCAAACTGGACAAGGAATCTTTGCTTTAAATGTGCCCGTCAGAGAGCTATT 242

QY 89 AsnMetLeuGlnGluIleMetGlnAsnSerIleAsnValValGluProValVal 108
 Db 243 AATATGTTGCAAGAGATAATGCAGATAATAGTATATAATGTTGGTAGAAGAACAGTAGTA 302
 QY 109 GluArgAsnAsnHisGlnThrGluLeuGluValProArgThrProArgThrProThrThr 128
 Db 303 GAAAGGGAATAACCAATCAATGAGTTGGAAGCTCCAAGAACCCCTCGGACACCTACCACT 362
 QY 129 ProGly 130
 Db 363 CCTGGG 368
 RESULT 14
 BU073369 669 bp mRNA linear EST 29-SEP-2003
 LOCUS BU073369 NIBB Mochii normalized Xenopus tailbud library Xenopus
 DEFINITION laevis cDNA clone XL109m21 5', mRNA sequence.
 ACCESSION BU073369
 VERSION BU073369.1 GI:17503558
 KEYWORDS EST.
 SOURCE Xenopus laevis (African clawed frog)
 ORGANISM Xenopus laevis
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
 Xenopodinae; Xenopus.
 REFERENCE 1 (bases 1 to 669)
 AUTHORS Kitayama,A., Terasaka,C., Mochii,M., Ueno,N., Shin-i,T. and
 Kohara,Y.
 TITLE Expressed genes in X. laevis embryo
 JOURNAL Unpublished (2001)
 COMMENT Contact: Tadasu Shin-i
 Center For Genetic Resource Information
 National Institute of Genetics
 1111 Yata, Mishima, Shizuoka 411-8540, Japan
 Tel: 81-559-81-6856
 Fax: 81-559-81-6855
 Email: tshini@genes.nig.ac.jp
 The information of this clone is available through the following
 URL:
 http://xenopus.nibb.ac.jp.
 FEATURES
 source
 1. 669
 Location/Qualifiers
 /organism="Xenopus laevis"
 /mol_type="mRNA"
 /db_xref="taxon:8355"
 /clone="XL109m21"
 /tissue_type="whole embryo"
 /dev_stage="stage 25"
 /clone_lib="NIBB Mochii normalized Xenopus tailbud
 library"

ORIGIN
 Alignment Scores:
 Pred. No.: 6,99e-68 Length: 669
 Score: 610.50 Matches: 115
 Percent Similarity: 96.12% Conservative: 9
 Best Local Similarity: 89.15% Mismatches: 4
 Query Match: 87.72% Indels: 1
 DB: 12 Gaps: 1
 US-09-757-415A-1_COPY_11_140 (1-130) x BU073369 (1-669)

QY 1 AspThrValProAspAsnHisArgAsnLysPheLysValIleAsnValAspAspGly 20
 Db 284 GAAACTATCCAGATAACCAACAAACAGATTAAAGTTATTATGTTGATGATGATGT 343
 QY 21 AsnGluLeuGlySerGlyIleMetGluLeuThrAspThrGluLeuIleuTyThrArg 40
 Db 344 AATGAGCTTGGCTCTGGCATTAATGGAATTAACCGAAATGATCTCATCTATACACTGC 403
 QY 41 LysArgAspSerValLysTrpHisIstYrLeuCysLeuArgArgTyGlyTyAspSerAsn 60
 Db 404 AACCGAGATTCTCTCAGATGGCTTACCTCTGCTTCGACGCTATGATGATGATCAAC 463

QY 61 LeuPheSerPheGluSerGlyArgArgCysGlnThrGlyGlnGlyIlePheAlaPheLys 80
 Db 464 CTTTTCCTTTTGAAAGTGGACGAAGTGTGAGACTGGACAGGAATATTGCTTTTAAA 523
 QY 81 CysAlaArgAlaGluGluLeuPheAsnMetLeuGlnGluIleMetGlnAsnSerIle 100
 Db 524 TGTGCCCGCGCAGAGAGCTCTTCAATATGCTTCAAGAGATCATGCAAAATAATAGCAT 583
 QY 101 AsnValValGluProValValGluArgAsnAsnHisGlnThrGluLeuGluValPro 120
 Db 584 AGCGTTGTGGAGGACCAAGTAGTTGAGCCA--AATCCCCAGACTGAGTTGATGTTCCA 640
 QY 121 ArgThrProArgThrProThrThrPro 129
 Db 641 AGAACCCCTCGCACACCAACCAACCCCT 667
 RESULT 15
 AL866682
 LOCUS AL866682 639 bp mRNA linear EST 03-DEC-2003
 DEFINITION AL866682 XGC-egg Silurana tropicalis cDNA clone TEGg126901 5', mRNA
 sequence.
 ACCESSION AL866682
 VERSION AL866682.2 GI:38662078
 KEYWORDS EST.
 SOURCE Silurana tropicalis (western clawed frog)
 ORGANISM Silurana tropicalis
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
 Xenopodinae; Silurana.
 REFERENCE 1 (bases 1 to 639)
 AUTHORS Croning,M.D.R., Ashurst,J.L., Taylor,R., Zorn,A.M. and Rogers,J.
 TITLE Sanger Xenopus tropicalis EST project 2001 (11_2003)
 JOURNAL Unpublished (2003)
 COMMENT On Sep 15, 2002 this sequence version replaced gi:22886947.
 Contact: Taylor R
 Sanger Institute
 Hinxton, Cambridgeshire, CB10 1SA, UK
 Email: trop@sanger.ac.uk
 Sanger Xenopus tropicalis EST project 2001
 TROPICALIS_SEQUENCE ID: TEGg126901.plkasp6
 Sequencing primer: SP6
 This sequence is from a Xenopus Gene Collection (XGC) library
 constructed by Aaron M. Zorn.
 cDNA was oligo dT primed from 5ug of poly A+ RNA from egg.
 EcoRI-NotI cut cDNA was then ligated into pCS107 with EcoRI at the
 5' end and NotI at the 3' end.
 Vector: pCS107; Site1: EcoRI; Site2: NotI
 Host: Escherichia coli XL1-blue.
 FEATURES
 Location/Qualifiers
 1. 639
 /organism="Silurana tropicalis"
 /mol_type="mRNA"
 /db_xref="taxon:8364"
 /clone="TEGg126901"
 /dev_stage="egg"
 /lab_host="Escherichia coli XL1-blue"
 /clone_lib="XGC-egg"
 /note="Vector: pCS107; Site 1: EcoRI; Site 2: NotI; cDNA
 was oligo dT primed from 5ug of poly A+ RNA from egg.
 EcoRI-NotI cut cDNA was then ligated into pCS107 with
 EcoRI at the 5' end and NotI at the 3' end."

ORIGIN
 Alignment Scores:
 Pred. No.: 8,77e-68 Length: 639
 Score: 609.50 Matches: 118
 Percent Similarity: 95.42% Conservative: 7
 Best Local Similarity: 90.08% Mismatches: 4
 Query Match: 87.57% Indels: 2
 DB: 9 Gaps: 1
 US-09-757-415A-1_COPY_11_140 (1-130) x AL866682 (1-639)

Qy	1	AspThrValProAspAsnHisArgAsnLysPheLysValIleAsnValAspAspGly	20
Db	225	GAACCTATCCCGATACCAACAACAGATTTAAGGTTATTACGTTGATGATGCT	284
Qy	21	AsnGluLeuGlySerGlyIleMetGluLeuThrAspThrGluLeuIleLeuIleValThrArg	40
Db	285	AATGACCTTGGCTCTGGCATATATGGAATTGCAGAAATGAGCTAATCTTATACACTCGT	344
Qy	41	LysArgAspSerValLysTrpHisTyrLeuCysLeuArgArgTyrGlyTyrAspSerAsn	60
Db	345	PAGCGAGATTCTGTCAGGTGGCCTTACCTCTGCTTGCAGCGCTATGCATACGATTCAAAC	404
Qy	61	LeuPheSerPheGluSerGlyArgArgCysGlnThrGlyGlnGlyIlePheAlaPheLys	80
Db	405	CTATTTTCTTTTCAAAGTGGACGAAGGTGTCAGACTGGACAAGGAATATTTGCTTTTAAG	464
Qy	81	CysAlaArgAlaGluLeuLeuPheAsnMetLeuGlnGluIleMetGlnAsnAsnSerIle	100
Db	465	TGTGCTCTGGTCGAGAGAGCTCTTCAATATGCTTCAAGAGATCATGCAGAACANTAGCAT	524
Qy	101	AsnValValGluGluProValValGluArgAsnAsnHisGln-ThrGluLeuGluValPro	120
Db	525	AGCGTTGTGGAGGAGCCAGTAGTTGAACGA--AATCCCAAGACTGAGTTGGATGTTCC	581
Qy	120	oArgThrProArgThrProThrThrProGly	130
Db	582	AAGAACCCTTCGCACACCCACACCCCTGGG	612

Search completed: August 9, 2004, 00:15:08
Job time : 3617.64 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 8, 2004, 21:10:45 ; Search time 99.2105 Seconds
(without alignments)

727.177 Million cell updates/sec

Title: US-09-757-415a-1_COPY_11_140

Perfect score: 696

Sequence: 1 DTVPDNRHNRKFKVINVDG.....NNHQTELEVPRTPRTPPG 130

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-O=/cgn2_1/USPTO spool/US09757415/runat_02082004_101232_6719/app_query.fasta_1.526
-DB=Issued Patents NA -QFMT=fastp -SUFFIX=p2n.rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=BITS -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09757415 @CGN 1.1.93 @runat_02082004_101232_6719 -NCPUS=6 -ICPU=3
-NO WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued Patents NA:
1: /cgn2_6/ptodata/2/ina/5A COMB.seq:
2: /cgn2_6/ptodata/2/ina/5B COMB.seq:
3: /cgn2_6/ptodata/2/ina/6A COMB.seq:
4: /cgn2_6/ptodata/2/ina/6B COMB.seq:
5: /cgn2_6/ptodata/2/ina/PCTUS COMB.seq:
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	143.5	20.6	1446	3	US-08-787-091-1
2	137.5	19.8	1146	4	US-09-620-312D-854
3	133.5	19.2	914	4	US-09-620-312D-855
4	124.5	17.9	2735	4	US-09-620-312D-842
5	83	11.9	1813	4	US-09-453-702B-91
6	82.5	11.9	415	4	US-09-621-976-12526
7	78	11.2	5828	4	US-09-023-655-1386
8	78	11.2	6152	1	US-08-557-139-1
9	77.5	11.1	698	3	US-08-896-164-44
10	75	10.8	1600	2	US-08-531-439B-3
11	75	10.8	1600	4	US-09-023-655-1505
12	74.5	10.7	2209	4	US-09-489-847-16

13	74	10.6	1664976	4	US-08-916-421B-1
14	74	10.6	4403765	3	US-09-103-840A-2
15	74	10.6	4411529	3	US-09-103-840A-1
16	73.5	10.6	1299	4	US-09-891-641-49
17	72.5	10.4	1626	4	US-09-134-001C-1644
18	72	10.3	2313	4	US-09-370-838-157
19	72	10.3	15082	4	US-09-004-838-89
20	70	10.1	13059	4	US-08-956-171B-220
21	69.5	10.0	3569	4	US-09-130-491-5
22	69	9.9	4403765	3	US-09-103-840A-2
23	69	9.9	4411529	3	US-09-103-840A-1
24	68.5	9.8	1029	4	US-09-543-681A-959
25	68.5	9.8	4865	4	US-08-894-017-24
26	68.5	9.8	4865	4	US-09-456-474-24
27	68.5	9.8	5177	6	5352450-1
28	68	9.8	825	4	US-09-134-001C-1594
29	68	9.8	4088	2	US-08-317-310A-1
30	68	9.8	4088	5	PCT-US95-13041-1
31	67.5	9.7	1029	4	US-09-489-039A-6921
32	67.5	9.7	2931	4	US-09-623-824-1
33	67.5	9.7	3815	4	US-08-936-165A-196
34	67	9.6	435	4	US-09-134-000C-1030
35	67	9.6	1394	4	US-09-215-694-25
36	67	9.6	1479	4	US-09-134-001C-1058
37	67	9.6	33000	4	US-09-215-694-18
38	67	9.6	1664976	4	US-08-916-421B-1
39	66.5	9.6	925	2	US-08-531-927B-8
40	66.5	9.6	1869	4	US-09-620-312D-704
41	66.5	9.6	6481	4	US-09-620-312D-1076
42	66.5	9.6	8876	4	US-08-961-527-170
43	66.5	9.6	12793	4	US-09-004-838-124
44	66	9.5	1137	4	US-09-134-001C-2650
45	65.5	9.4	1238	4	US-09-761-962A-2

ALIGNMENTS

RESULT 1

US-08-787-091-1
; Sequence 1, Application US/08787091
; Patent No. 6100386
; GENERAL INFORMATION:
; APPLICANT: Carpino, Nicholas A.
; APPLICANT: Kobayashi, Ryuji
; APPLICANT: Wisniewski, David G.
; APPLICANT: Strife, Annabel O'C.
; APPLICANT: Clarkson, Bayard D.
; TITLE OF INVENTION: A No. 6100386el Human Gene/Protein Involved in
; TITLE OF INVENTION: Chronic Myelogenous Leukemia
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: MA
; COUNTRY: US
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/787,091
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/030,418
; FILING DATE: 01-NOV-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Granahan, Patricia
; REGISTRATION NUMBER: 32,227
; REFERENCE/DOCKET NUMBER: CSHL96-05pa

TELECOMMUNICATION INFORMATION:

TELEPHONE: 617-861-6240
 TELEFAX: 617-861-9540
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1446 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 1..1443

US-08-787-091-1

Alignment Scores:

Pred. NO.: 3.13e-10 Length: 1446
 Score: 143.50 Matches: 33
 Percent Similarity: 47.32% Conservatives: 20
 Best Local Similarity: 29.46% Mismatches: 50
 Query Match: 20.62% Indels: 9
 Gaps: 3

US-09-757-415A-1_COPY_11_140 (1-130) x US-08-787-091-1 (1-1446)

QY 24 GlySerGlyIleMetGluLeuThrAspThrGluLeuIleLeuTyThrArgLysArgAsp 43
 Db 514 GGCTCTACGTGCTGAGGGTGGAGGCTGAAGCTGACTCTCTGACCGTGGGGCCCGAG 573
 QY 44 Ser-----ValysTrpHisTyLeuCysLeuArgArgTyGlyTyr 57
 Db 574 AGTCAGATAGTGGAGCCACTCTGTCTGCTGCCCTACACTCTGTTGGTGGCTATGCGCG 633
 QY 58 AspSerAsnLeuPheSerPheGluSerGlyArgCysGlnThrGlyGlnGlyIlePhe 77
 Db 634 GACAAGGTGATGTTCTTTTCGAGGCGCGCGCGCTGCGCCCTCAGGCGCTGGAACCTTC 693
 QY 78 AlaPheLysCysAlaArgAlaGluLeuPheAsnMetLeuGlnGluIleMetGlnAsn 97
 Db 694 ACCTTCAGCAGCGCACAGGGAATGACATCTTCAGGCAGTTGAGACTGCCATCCACCGG 753
 QY 98 AsnSerIleAsnValValGlu-----GluProValValGluArgAsnAsnHisGln 114
 Db 754 CAGAAGCCCGAGGAAGCGCGGACAGGCGCAGATGTTCTTCAGAGCTGACTCCCATGAA 813
 QY 115 ThrGluLeuGluValProArgThrProArgThrPro 126
 Db 814 GGGGAGGTGGCAGAGGGGAAGTTGCTTCCCCACCT 849

RESULT 2

US-09-620-312D-854
 ; Sequence 854, Application US/09620312D
 ; Patent No. 6569662

GENERAL INFORMATION:

APPLICANT: Tang, Y. Tom
 APPLICANT: Liu, Chenghua
 APPLICANT: Asundi, Vinod
 APPLICANT: Zhang, Jie
 APPLICANT: Ren, Feiyan
 APPLICANT: Chen, Rui-hong
 APPLICANT: Zhao, Qing A.
 APPLICANT: Wehrman, Tom
 APPLICANT: Xue, Aidong J.
 APPLICANT: Wang, Yonghong
 APPLICANT: Wang, Jian-Rui
 APPLICANT: Zhou, Ping
 APPLICANT: Ma, Yungqing
 APPLICANT: Wang, Dunrui
 APPLICANT: Wang, Zhiwei
 APPLICANT: John Tillinghast
 APPLICANT: Drmanac, Radoje T.
 TITLE OF INVENTION: No. 6569662el Nucleic Acids and
 TITLE OF INVENTION: Polypeptides

FILE REFERENCE: 784CIP2B

CURRENT APPLICATION NUMBER: US/09/620,312D
 CURRENT FILING DATE: 2000-07-19
 PRIOR APPLICATION NUMBER: 09/552,317
 PRIOR FILING DATE: 2000-04-25
 PRIOR APPLICATION NUMBER: 09/488,725
 PRIOR FILING DATE: 2000-01-21
 NUMBER OF SEQ ID NOS: 1105
 SOFTWARE: pt_FL_genes Version 1.0
 SEQ ID NO 854

LENGTH: 1146

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: CDS

LOCATION: (109) .. (1122)

US-09-620-312D-854

Alignment Scores:

Pred. NO.: 1.55e-09 Length: 1146
 Score: 137.50 Matches: 34
 Percent Similarity: 49.07% Conservatives: 19
 Best Local Similarity: 31.48% Mismatches: 42
 Query Match: 19.76% Indels: 13
 Gaps: 4

US-09-757-415A-1_COPY_11_140 (1-130) x US-09-620-312D-854 (1-1146)

QY 24 GlySerGlyIleMetGluLeuThrAspThrGluLeuIleLeuTyThrArgLysArgAsp 43
 Db 565 GGTGAATCCACAATGACATCACTCATGAAATATCTCTCTGGATATCCACAATGCC 624
 QY 44 SerValys-----TrpHisTyLeuCysLeuArgArgTyGlyTyrAspSerAsn 60
 Db 625 AAGTCAAACTGGTGATGGCTCTCAGCTCACTGAGGAGATACGGTCCGGACTCAACG 684
 QY 61 LeuPheSerPheGluSerGlyArgCysGlnThrGlyGlnGlyIlePheAlaPheLys 80
 Db 685 TGGTTCAGTTGAGTCAGGAGATGTCACACAGGAGAGGACTATTTCATTTTCAA 744
 QY 81 CysAlaArgAlaGluLeuPheAsnMetLeuGlnGluIleMetGlnAsnAsnSerIle 100
 Db 745 ACAAGGGAAGGAGAA-----ATGATCTATCAGAAAGTTTCTTCTCGCACACTG 792
 QY 101 AsnValValGluGlu-----ProValValGluArgAsnAsnHisGln 114
 Db 793 GCCATAGCTGAGCAACATGAAGATTAACTAGTAATGGAACAGAGGCCCGCTTCAG 852
 QY 115 ThrGluLeuGluValProArgThr 122
 Db 853 ACAAGCTTGACTGAACCAATGACA 876

RESULT 3

US-09-620-312D-855

; Sequence 855, Application US/09620312D

; Patent No. 6569662

GENERAL INFORMATION:

APPLICANT: Tang, Y. Tom
 APPLICANT: Liu, Chenghua
 APPLICANT: Asundi, Vinod
 APPLICANT: Zhang, Jie
 APPLICANT: Ren, Feiyan
 APPLICANT: Chen, Rui-hong
 APPLICANT: Zhao, Qing A.
 APPLICANT: Wehrman, Tom
 APPLICANT: Xue, Aidong J.
 APPLICANT: Wang, Yonghong
 APPLICANT: Wang, Jian-Rui
 APPLICANT: Zhou, Ping
 APPLICANT: Ma, Yungqing
 APPLICANT: Wang, Dunrui
 APPLICANT: Wang, Zhiwei
 APPLICANT: John Tillinghast

```

; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. 6569662el Nucleic Acids and
; FILE REFERENCE: 784CIP2B
; CURRENT APPLICATION NUMBER: US/09/620,312D
; CURRENT FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1105
; SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 855
; LENGTH: 914
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (109)..(867)
US-09-620-312D-855

Alignment Scores:
Pred. No.: 4,04e-09 Length: 914
Score: 133.50 Matches: 33
Percent Similarity: 48.62% Conservative: 20
Best Local Similarity: 30.28% Mismatches: 43
Query Match: 19.18% Indels: 13
DB: 4 Gaps: 3

US-09-757-415a-1_copy_11_140 (1-130) x US-09-620-312D-855 (1-914)
QY 24 GlySerGlyLeuMetGluLeuThrAspThrGluLeuLeuLeuLeuTyrThrArgLysArgAsp 43
DB 565 GGTGAATGACCAATGCAGATCACTCATGAAATATCTATCTCTGGGATATCCCAATGCC 624
QY 44 SerValLys-----TrpHisTyrLeuCysLeuArgArgTyrGlyTyrAspSerAsn 60
DB 625 AAGGTCAACTGGTGATGGCCCTCTCAGCTCACTGAGGAGATCGGTGGGACTCAACG 684
QY 61 LeuPheSerPheGluSerGlyArgArgCysGlnThrGlyGlnGlyLeuPheAlaPheLys 80
DB 685 TGGTTTCACGTTTCAGTCAGGAGAAATGTGTGACACAGGAGAGACTATTCATTTCAA 744
QY 81 CysAlaArgAlaGluLeuPheAsnMetLeuGlnGluLeuMetGlnAsnSerIle 100
DB 745 ACAAGGGAAGGAGAA-----ATGATCTATCAGAAGTTTCATCTCGGACACG 792
QY 101 AsnValValGluGluProValValGluArgAsnAsnHisGlnThrGluLeuGluValPro 120
DB 793 GCATAGCTAGGCAA-----CATGAAAGATTAATGCTAGAAATGGAA 834
QY 121 ArgThrProArgThrProThrPro 129
DB 835 CAGAAGGCCCGGTAAAGGCCCTTCTCT 861

RESULT 4
US-09-620-312D-842
; Sequence 842, Application US/09620312D
; Patent No. 6569662
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyang
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yongshong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yunqing

; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: John Tillinghast
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. 6569662el Nucleic Acids and
; FILE REFERENCE: 784CIP2B
; CURRENT APPLICATION NUMBER: US/09/620,312D
; CURRENT FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1105
; SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 842
; LENGTH: 2735
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (334)..(1314)
US-09-620-312D-842

Alignment Scores:
Pred. No.: 3,77e-07 Length: 2735
Score: 124.50 Matches: 31
Percent Similarity: 46.73% Conservative: 19
Best Local Similarity: 28.97% Mismatches: 48
Query Match: 17.89% Indels: 9
DB: 4 Gaps: 3

US-09-757-415a-1_copy_11_140 (1-130) x US-09-620-312D-842 (1-2735)
QY 28 MetGluLeuThrAspThrGluLeuLeuLeuTyrThrArgLysArgAspSerValLys--- 46
DB 802 CTGCAGATCAACCCAGAGAACATCTCTCTGGGACATCCACACACCCCGTGTGAAGTC 861
QY 47 -----TrpHisTyrLeuCysLeuArgArgTyrGlyTyrAspSerAsnLeuPheSerPhe 64
DB 862 GTCTGTCGGCCCTCTGCTCACTGCGCGCTATGGCCGGGATGCCACACGCTTTACCTTC 921
QY 65 GluSerGlyArgArgCysGlnThrGlyGlnGlyLeuPheAlaPheLysCysAlaArgAla 84
DB 922 GAGGTGCGCGGATGTGTGATGCTGGGAGAGGACTCTATACCTTCCAGACACACAGAGGG 981
QY 85 GluGluLeuPheAsnMetLeuGlnGluLeuMet-----GlnAsnAsnSer 99
DB 982 GAGCAGATTTACCCAGCGGTCCACAGTCCACCTGGCCATCGCAGACACACAGCGG 1041
QY 100 IleAsnValValGluGluProValValGluArgAsnAsnHisGlnThrGlu----LeuGlu 118
DB 1042 GTCTGCTGGAAATGGAGAGAACTGAGGTGCTGAAACAGGGCAGCGAATTTACTCG 1101
QY 119 ValProArgThrProArgThr 125
DB 1102 TATCCCTGCACACCCACGACC 1122

RESULT 5
US-09-453-702B-91/C
; Sequence 91, Application US/09453702B
; Patent No. 6365723
; GENERAL INFORMATION:
; APPLICANT: Blattner, Frederick R.
; APPLICANT: Burland, Valerie
; APPLICANT: Perna, Nicole T.
; APPLICANT: Plunkett, Guy
; APPLICANT: Welch, Rod
; TITLE OF INVENTION: No. 6365723el Sequences of E. coli O157
; NUMBER OF SEQUENCES: 265
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Quarles & Brady
; STREET: 1 South Pinckney Street
```

```
; Sequence 12526, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621.976
; NUMBER OF SEQ ID NOS: 2000-07-21
; FILING DATE: 2000-07-21
; SOFTWARE: Patent.pm
; SEQ ID NO 12526
; LENGTH: 415
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-621-976-12526

Alignment Scores:
Pred. No.: 0.0186 Length: 415
Score: 82.50 Matches: 20
Percent Similarity: 52.08% Conservative: 5
Best Local Similarity: 41.67% Mismatches: 20
Query Match: 11.85% Indels: 3
DB: 4 Gaps: 1

US-09-757-415A-1_COPY_11_140 (1-130) x US-09-621-976-12526 (1-415)
QY 24 GlySerGlyIleMetGluLeuThrAspThrGluLeuIleLeuTyrThrArgLysArgAsp 43
Db 268 GGTGAATGCACATGCAGATCATCATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 327
QY 44 SerValLys-----TTPHisTyrLeuCysLeuArgArgTyrGlyTyrAspSerAsn 60
Db 328 AAGTCAACTGGTGATGTGGCTCTCAGTCACTGAGGAGATACGTCGGGACTCAACG 387
QY 61 LeuPheSerPheGluSerGlyArg 68
Db 388 TGGTTCAGTTTGGTTCAGGAAGA 411

RESULT 7
US-09-023-655-1386
; Sequence 1386, Application US/09023655
; Patent No. 6607879
; GENERAL INFORMATION:
; APPLICANT: Cocks, Benjamin G.
; APPLICANT: Susan G. Stuart
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
; TITLE OF INVENTION: EXPRESSION
; NUMBER OF SEQUENCES: 1508
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 1174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/023.655
; FILING DATE: HEREWITH
; CLASSIFICATION:
; PRIOR APPLICATION NUMBER:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
```

REGISTRATION NUMBER: 37,071
REFERENCE/DOCKET NUMBER: PA-0001 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 867-0123
TELEFAX: (212) 878-9655
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 5828 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GENBANK
CLONE: G386256
US-09-023-655-1386

Alignment Scores:
Pred. No.: 4 Length: 5828
Score: 78.00 Matches: 21
Percent Similarity: 40.00% Conservative: 7
Best Local Similarity: 30.00% Mismatches: 20
Query Match: 11.21% Indels: 22
Gaps: 1

US-09-757-415A-1_COPY_11_140 (1-130) x US-09-023-655-1386 (1-5828)

QY 52 LeuArgArgTyrGlyTyrAspSerAsnLeuPheSerPheGluSerGlyArgGlyCysGln 71
DB 1651 ATCAGCGCTGTGGCCACTCGAAACTTCTTTCATCGAGTGGCGGTTCTGCCGTG 1710
QY 72 ThrGlyGlnGlyIlePheAlaPheLysCysAlaArgAlaGluLeuPheAsnMetLeu 91
DB 1711 ACGGGGCGGGGAGTTCTGGATCAG----- 1737
QY 92 GlnGluMetGlnAsnAsnSerIleAsnValValGluGluProValValGluArgAsn 111
DB 1738 -----GTGATGACTCTGTGGTGGCCAGAAC 1764
QY 112 AsnHisGlnThrGluLeuGluValProArg 121
DB 1765 ATGCACGAGACCATCTCGAGGCCATGCGG 1794

RESULT 8
US-08-557-139-1
Sequence 1, Application US/08557139
Patent No. 5827730
GENERAL INFORMATION:
APPLICANT: Pedersen, Oluf
APPLICANT: Bjorkbak, Christian
APPLICANT: Frederiksen, Kathrine A.
TITLE OF INVENTION: MUTANT DNA ENCODING INSULIN RECEPTOR
NUMBER OF SEQUENCES: 1
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 58277300 No. 5827730disk of No. 5827730th America
STREET: 405 Lexington Avenue
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10174
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/557,139
FILING DATE: 12-FEB-1996
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Lambiris, Elias J.
REGISTRATION NUMBER: 33,728
REFERENCE/DOCKET NUMBER: 4041.204-US

TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 867-0123
TELEFAX: (212) 878-9655
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 6152 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 581...4309
US-08-557-139-1

Alignment Scores:
Pred. No.: 4.34 Length: 6152
Score: 78.00 Matches: 21
Percent Similarity: 40.00% Conservative: 7
Best Local Similarity: 30.00% Mismatches: 20
Query Match: 11.21% Indels: 22
Gaps: 1

US-09-757-415A-1_COPY_11_140 (1-130) x US-08-557-139-1 (1-6152)

QY 52 LeuArgArgTyrGlyTyrAspSerAsnLeuPheSerPheGluSerGlyArgGlyCysGln 71
DB 1214 ATCAGCGCTGTGGCCACTCGAAACTTCTTTCATCGAGTGGCGGTTCTGCCGTG 1273
QY 72 ThrGlyGlnGlyIlePheAlaPheLysCysAlaArgAlaGluLeuPheAsnMetLeu 91
DB 1274 ACGGGGCGGGGAGTTCTGGATCAG----- 1300
QY 92 GlnGluMetGlnAsnAsnSerIleAsnValValGluGluProValValGluArgAsn 111
DB 1301 -----GTGATGACTCTGTGGTGGCCAGAAC 1327
QY 112 AsnHisGlnThrGluLeuGluValProArg 121
DB 1328 ATGCACGAGACCATCTCGAGGCCATGCGG 1357

RESULT 9
US-08-896-164-44
Sequence 44, Application US/08896164
Patent No. 6218521
GENERAL INFORMATION:
APPLICANT: OBATA, Yuichi
TITLE OF INVENTION: ISOLATED NUCLEIC ACID MOLECULES ASSOCIATED
TITLE OF INVENTION: WITH GASTRIC CANCER AND METHODS FOR
TITLE OF INVENTION: DIAGNOSING AND TREATING GASTRIC CANCER
NUMBER OF SEQUENCES: 87
CORRESPONDENCE ADDRESS:
ADDRESSEE: Felfe & Lynch
STREET: 805 Third Avenue
CITY: New York City
STATE: New York
COUNTRY: USA
ZIP: 10022
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.44mb
COMPUTER: IBM PS/2
OPERATING SYSTEM: PC-DOS
SOFTWARE: Wordperfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/896,164
FILING DATE: July 17, 1997
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: No. 6218521man D. Hanson
REGISTRATION NUMBER: 30,946
REFERENCE/DOCKET NUMBER: LUD 5499 - JEL/NDH/SLH
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 688-9200

TELEFAX: (212) 838-3884
INFORMATION FOR SEQ ID NO: 44:
SEQUENCE CHARACTERISTICS:
LENGTH: 698 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-896-164-44

Alignment Scores:
Pred. No.: 0.203 Length: 698
Score: 77.50 Matches: 27
Percent Similarity: 43.85% Conservative: 30
Best Local Similarity: 20.77% Mismatches: 50
Query Match: 11.14% Indels: 23
DB: 3 Gaps: 6

US-09-757-415A-1_COPY_11_140 (1-130) x US-08-896-164-44 (1-698)

QY 2 ThrValProAspAsnHisArgAsnLysPheLysValIleAsnValAspAspGlyAsn 21
Db 285 TCCATTCCAAAAAAGGAAATCTGTAGTTACCAATGAGAATGAT----- 338
QY 22 GluLeuGlySerGlyIleMetGluLeuThr-----AspThrGluLeuIleLeuTyr 38
Db 339 -----ACTCTCTAGAGAACTCTCAAAATATGTAGACATCATATATTGCGCCTT 389
QY 39 ThrArgLysArgAspSerValLysTrpHisTyr-----LeuCysLeuArgArgTyrGly 56
Db 390 ACTCGAAATCGAGGACAAAGAGATGTACACTGTCTCCACTGTGTGGGAAACAGTTAAT 449
QY 57 TyrAspSerAsnLeuPheSerPheGluSerGlyArgArgCysGlnThrGlyGlnGlyIle 76
Db 450 GAAAGTCTTACCTCATTCTCCACCAG-----AGGACCACACTCGAGAAAAA--- 497
QY 77 PheAlaPheLysCysAlaArgAlaGluGluLeuPheAsnMetLeuGlnGluIleMetGln 96
Db 498 ---CCCTATGACTGTNTCACTCTGGGAAAGCTTCAATCATNAACAACCTCAATAA 554
QY 97 AsnAsnSerIleAsnValGluGluProVal-----Val 108
Db 555 CATGAGCGAATTCNTACAGAGAGAAACCTTATCTCTGTTCTCAGTGTGGAAAAAACTTC 614
QY 109 GluArgAsnAsnHisGlnThrGluLeuGlu 118
Db 615 CGTCNGAATTCATCGGAGTCTCTCGAA 644

RESULT 10

US-08-531-439B-3
Sequence 3, Application US/08531439B
Patent No. 5981702

GENERAL INFORMATION:

APPLICANT: Zhang, Hui
TITLE OF INVENTION: Cyclin/CDK Associated Proteins,
and Uses Related Thereto
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESS: Foley, Hoag & Eliot
STREET: One Post Office Square
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII (text)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/531,439B
FILING DATE: 21-SEP-1995
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Vincent, Matthew P.
REGISTRATION NUMBER: 36,709
REFERENCE/DOCKET NUMBER: CSV-006.01
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 832-1000
TELEFAX: (617) 832-7000
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1600 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 148..1452
US-08-531-439B-3

Alignment Scores:
Pred. No.: 1.56 Length: 1600
Score: 75.00 Matches: 40
Percent Similarity: 36.90% Conservative: 22
Best Local Similarity: 23.81% Mismatches: 58
Query Match: 10.78% Indels: 48
DB: 2 Gaps: 8

US-09-757-415A-1_COPY_11_140 (1-130) x US-08-531-439B-3 (1-1600)

QY 4 ProAspAsnHisArgAsnLysPheLysValIleAsnValAspAspGlyAsnGluLeu 23
Db 436 CCTAAGCTAAATCGGAGAACTTTCCAGGTGTTTCATGGGACTCTCTTCGGATGAGCTG 495
QY 24 GlySerGlyIleMetGlu---LeuThrAspThrGluLeuIle----- 36
Db 496 CTCITGGGAATCTTTCTCTGTCCTGCTGCTCCCTGAGCTGCTAAAGGTCTCTGGTGTGTGT 555
QY 37 -----LeuTyrThrArgLysArgAspSerValLysTrpHisTyrLeuCysLeuArgArg 54
Db 556 AAGAGGTGTATCGCTAGCGTCTGTAGTCTCTATGGCAGACCTTAGACCTTACA--- 612
QY 55 TyrGlyTyrAspSerAsnLeuPheSerPheGluSerGlyArgArgCysGlnThrGlyGln 74
Db 613 -----GGTAAATCTGCACCCGAGTGTACTGTCTGG-----TTGCTGTCTCAA 657
QY 75 GlyIlePheAlaPheLysCysAlaArgAla----- 84
Db 658 GGGGTGATTGCTTCCGCTGCCACGATCATTTATGGACCAACCATTTGGCTGAACATTC 717
QY 85 -----GluGluLeuPheAsnMetLeuGlnGluIle----- 94
Db 718 AGCCCTTTTGTGTACAGGACATGGACCTATCGAATCTAGTATAGAAAGTGTCCACCCCTC 777
QY 95 -----MetGlnAsnAsnSerIleAsnVal---ValGlu 104
Db 778 CACGGCATACTGTCTCAGTGTTCACAGTTCCAGAACTTAAGCTCGGAACCTGCGGCTTCG 837
QY 105 GluProValValGlu-----ArgAsnAsnHisGlnThrGluLeuGluValProArg 121
Db 838 GATCCCATTTGCAATACTCTCGCAAAAACTCAAAATTTAGTGGCACTTAACCTTCTCTGGG 897
QY 122 ThrProArgThrProThrPro 129
Db 898 TGTCTGATTCTCCATAATTTCCC 921

RESULT 11

US-09-023-655-1505
Sequence 1505, Application US/09023655
Patent No. 6607879
GENERAL INFORMATION:
APPLICANT: Cocks, Benjamin G.
APPLICANT: Susan G. Stuart
APPLICANT: Jeffrey J. Seilhamer

TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE

NUMBER OF SEQUENCES: 1508

CORRESPONDENCE ADDRESS:

ADDRESSEE: INCYTE PHARMACEUTICALS, INC.

STREET: 3174 PORTER DRIVE

CITY: PALO ALTO

STATE: CALIFORNIA

COUNTRY: USA

ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/023,655

FILING DATE: HEREWITH

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Zeller, Karen J.

REGISTRATION NUMBER: 37,071

REFERENCE/DOCKET NUMBER: PA-0001 US

TELECOMMUNICATION INFORMATION:

TELEPHONE: (650) 855-0555

TELEFAX: (650) 845-4166

INFORMATION FOR SEQ ID NO: 1505:

SEQUENCE CHARACTERISTICS:

LENGTH: 1600 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

IMMEDIATE SOURCE:

LIBRARY: GENBANK

CLONE: g995825

US-09-023-655-1505

Alignment Scores:
Pred. No.: 1,56 Length: 1600
Score: 75.00 Matches: 40
Percent Similarity: 36.90% Conservative: 22
Best Local Similarity: 23.81% Mismatches: 58
Query Match: 10.78% Indels: 48
DB: 4 Gaps: 8

US-09-757-415A-1_COPY_11_140 (1-130) x US-09-023-655-1505 (1-1600)

QY 4 ProAspAsnHisArgAsnLysPheLysValIleAsnValAspAspGlyAsnGluLeu 23
DB 436 CCTAAGCTAAATCGGGAGAACTTCCAGGTGTTTCATGGGACTCTCTCCGGAAGCTG 495
QY 24 GlySerGlyIleMetGlu---LeuThrAspThrGluLeuIle----- 36
DB 496 CTCTTGGGAATCTTTCTGCTGCTGCTGCTCCCTGAGCTGCTAAAGGCTCTGCTGTTGT 555
QY 37 -----LeuThrThrArgLysArgAspSerValIysTrpHisTyrLeuCysLeuArg 54
DB 556 AAGAGGTGGTATCGCTAGCGTCTGTATGAGTCTCTATGGCAGACCTTAGACCTTACA--- 612
QY 55 TyrGlyTyrAspSerAsnLeuPheSerPheGluSerGlyArgArgCysGlnThrGlyGln 74
DB 613 -----GGTAAATCTCCACCGGATGTGACTGTGCG-----TTGCTGTCTCAA 657
QY 75 GlyIlePheAlaPheLysCysAlaArgAla----- 84
DB 658 GGGGTGATTGCTTCGGTGGCCACCATATTATGACCAACCATTTGGGTGAACATTC 717
QY 85 -----GluGluLeuPheAsnMetLeuGlnGluIle----- 94

DB 718 AGCCCTTTTCGTGTACAGGACATGGACCTATCGAACTCAGTTATAGAGTGTCCACCTC 777
QY 95 -----MetGlnAsnAsnSerIleAsnVal---ValGlu 104
DB 778 CACGGCATACTGTCTCAGTGTCCAGTTCGAGAATCTAAGCTCGGAACCTGGCGCTTTCG 837
QY 105 GluProValValGlu-----ArgAsnAsnHisGlnThrGluLeuGluValProArg 121
DB 838 GATCCCATTTGTCAATACTCTCGCAAAAACTCAAATTTAGTGGACATTAACTTCTCTGGG 897
QY 122 ThrProArgThrProThrThrPro 129
DB 898 TGTCTGGATTCCCTAAATTTCCC 921
RESULT 12
US-09-489-847-16
; Sequence 16, Application US/09489847
; Patent No. 6476195
; GENERAL INFORMATION:
; APPLICANT: Rosen et al
; TITLE OF INVENTION: 98 Human Secreted Proteins
; FILE REFERENCE: PZ031P1
; CURRENT APPLICATION NUMBER: US/09/489,847
; CURRENT FILING DATE: 2000-01-24
; EARLIER APPLICATION NUMBER: PCT/US99/17130
; EARLIER FILING DATE: 1999-07-29
; EARLIER APPLICATION NUMBER: 60/094,657
; EARLIER FILING DATE: 1998-07-30
; EARLIER APPLICATION NUMBER: 60/095,486
; EARLIER FILING DATE: 1998-08-05
; EARLIER APPLICATION NUMBER: 60/096,319
; EARLIER FILING DATE: 1998-08-12
; EARLIER APPLICATION NUMBER: 60/095,454
; EARLIER FILING DATE: 1998-08-06
; EARLIER APPLICATION NUMBER: 60/095,455
; EARLIER FILING DATE: 1998-08-06
; NUMBER OF SEQ ID NOS: 376
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 16
; LENGTH: 2209
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-489-847-16
Alignment Scores:
Pred. No.: 2,95 Length: 2209
Score: 74.50 Matches: 30
Percent Similarity: 39.50% Conservative: 17
Best Local Similarity: 25.21% Mismatches: 53
Query Match: 10.70% Indels: 19
DB: 4 Gaps: 4
US-09-757-415A-1_COPY_11_140 (1-130) x US-09-489-847-16 (1-2209)
QY 26 GlyIleMetGluLeuThr---AspThrGluLeuIleLeuTyrThrArgLysArgAsp--- 43
DB 1713 GCGGTGTTTTCAGGTGACCCCTGACACAAAGGCTTACCTGAGGACCCCAACTGG 1772
QY 44 -----SerValIysTrpHisTyrLeuCysLeuArgTyr 55
DB 1773 GACCGGGGCTGCCATCCCTCACCCTCTGTCTCGAACATCAGCGTCCAGACACAG 1832
QY 56 GlyTyrAspSerAsnLeuPheSerPheGluSerGlyArgArgCysGlnThrGlyGlnGly 75
DB 1833 GTGGCTGCTGCTGCTGCTTTTAAAGGAGCGGCGGTGCTGCCAGACAGGCGCGCA 1892
QY 76 -----IlePheAlaPheLysCysAlaArgAlaGluLeuPheAsnMetLeuGlnGlu 93
DB 1893 TTCATGATCATCCAGGACGCGGCGGCCCGGTGAGGAGATCTTCAGCTGACCGGAT 1952
QY 94 IleMetGlnAsnAsnSer-----IleAsnValValGluGluPro 106
DB 1953 GTGCTCCCAAGCAAGCTTCCACCATCACAGCTTCTGGGTCAACATCTTAAATGMAGC 2012

QY 107 ValValGluArgAsnAsnHisGlnThrGluLeuValProArgThrProArgThr 125
US-08-916-421B-1
; Sequence 1, Application US/08916421B
; Patent No. 6503729
; GENERAL INFORMATION:
; APPLICANT: Buit et al.
; TITLE OF INVENTION: Complete Genome Sequence of the Methanogenic Archaeon, Methanococcus
; Patent No. 6503729
; TITLE OF INVENTION: jannaschii
; FILE REFERENCE: PB275
; CURRENT APPLICATION NUMBER: US/08/916,421B
; CURRENT FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: US 60/024,428
; PRIOR FILING DATE: 1996-08-22
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1664976
; TYPE: DNA
; ORGANISM: Methanococcus jannaschii
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (28222)..(28222)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (28257)..(28258)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (84773)..(84773)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (84808)..(84808)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (84812)..(84812)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (98120)..(98120)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (98159)..(98159)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (98239)..(98239)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (98266)..(98266)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (98343)..(98343)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (103998)..(103998)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (148948)..(148948)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (163385)..(163385)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (191989)..(191989)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (191995)..(191995)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (231980)..(231980)
; OTHER INFORMATION: n equals a, t, c, or g
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (234187)..(234187)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (234220)..(234220)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (234814)..(234814)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (309398)..(309398)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (309418)..(309418)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (312837)..(312837)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (312993)..(312993)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (319226)..(319226)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (559167)..(559167)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (559241)..(559241)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (600992)..(600992)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (622708)..(622708)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (657081)..(657081)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (657203)..(657203)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (674435)..(674435)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (682442)..(682442)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (713652)..(713652)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (71684)..(71684)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (779455)..(779455)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (779676)..(779676)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (855539)..(855539)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (871619)..(871619)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1084830)..(1084830)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1096846)..(1096846)
; OTHER INFORMATION: n equals a, t, c, or g


```
; NAME/KEY: misc feature
; LOCATION: (1119881)..(1119881)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1130881)..(1130881)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1310988)..(1310988)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1313224)..(1313224)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1349473)..(1349473)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1349491)..(1349491)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1470091)..(1470091)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1603734)..(1603734)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1637998)..(1637998)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc feature
; LOCATION: (1664854)..(1664854)
; OTHER INFORMATION: n equals a, t, c, or g
; US-08-316-421B-1

Alignment Scores:
Pred. No.: 6.2e+04 Length: 1664976
Score: 74.00 Matches: 20
Percent Similarity: 45.16% Conservative: 8
Best Local Similarity: 32.26% Mismatches: 20
Query Match: 10.63% Indels: 14
DB: 4 Gaps: 2

US-09-757-415A-1_COPY_11_140 (1-130) x US-08-916-421B-1 (1-1664976)
QY 47 TrpHisTyrLeuCysLeuArgAgtTyrGlyTyrAspSerAsnLeuPheSerPheGluSer 66
Db 582557 TGGTGTATTATTATTAACACAGAGATATTTCTGAT---AATCTGCTTTCATTTTATAAC 582613
QY 67 GlyArgArgCysGlnThrGlyGlnGlyLePheAlaPheLysCysAlaArgAlaGluGlu 86
Db 582614 AATAAACAAAGTAATGCTCTATTAGGGG----- 582640
QY 87 LeuPheAsnMetLeuGlnGluLeMetGlnAsnAsnSerIleAsnValValGluGluPro 106
Db 582641 -----ACATGCTGGAGGAGATATATCAAAATCATCACTGAATAAATAACATCAGCC 582694
QY 107 ValVal 108
Db 582695 ACAGTA 582700

RESULT 14
US-09-103-840A-2/c
; Sequence 2, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4411529

; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4411529

; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4403765
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; OTHER INFORMATION: CDC 1551
; OTHER INFORMATION: "n" bases at various positions throughout the sequence
; OTHER INFORMATION: represent a, t, c or g
; US-09-103-840A-2

Alignment Scores:
Pred. No.: 2.43e+05 Length: 4403765
Score: 74.00 Matches: 33
Percent Similarity: 45.45% Conservative: 22
Best Local Similarity: 27.27% Mismatches: 49
Query Match: 10.63% Indels: 17
DB: 3 Gaps: 5

US-09-757-415A-1_COPY_11_140 (1-130) x US-09-103-840A-2 (1-4403765)
QY 16 ValAspAspAspGlyAsnGluLeuGlySerGlyIleMetGluLeuThrAspThrGluLeu 35
Db 2083125 GTCGACACTGACGGCAGCGAGGTGTCGAAGCATCATGATGATCCCTATTCACCGCCCTG 2083066
QY 36 Ile---LeuTyrThrArgLysArgAspSerValIysTrpHisTyrLeuCysLeuArg 54
Db 2083065 ATCGCCATCCACGGCTAAAGGCCAGCGACGTCAACGGCGCGCTGATC----- 2083018
QY 55 TyrGlyTyrAspSerAsnLeuPheSerPheGluSerGlyArgArgCysGlnThrGlyGln 74
Db 2083017 -----AACAGCGCGCAGCGGCTCCATCTACATCGTCAAGCCGAGCATGACCGTCCG 2082967
QY 75 GlyIlePheAlaPheLysCysAlaArgAlaGluGluLeuPheAsnMetLeuGlnGluLeu 94
Db 2082966 GCCGAGGTGGCGTTTACCTGC-----GNACTGTTTCAGCGCGGTTCAAGATGTG 2082919
QY 95 Met-----GlnAsnAsn---SerIleAsnValValGluGluPro-ValValGluArg 110
Db 2082918 CTGGGGTTTCCGCAAAACACCATGAAGATCGGCATCATGAGGAGGAGGAGGAGGAGGAGG 2082859
QY 110 gAsnAsnHisGlnThrGluLeuGluValProArgThrProArgThrProThrThrProGl 130
Db 2082859 GTCAACCTCAAGCGGTGCATCAAGCTGCGCGGACCGCGTGTCTTTCATCAACACCGGG 2082799
QY 130 Y 130
Db 2082798 T 2082798

RESULT 15
US-09-103-840A-1/c
; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4411529
```

; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37Rv
US-09-103-840A-1

Alignment Scores:

Pred. No.:	2.44e+05	Length:	4411529
Score:	74.00	Matches:	33
Percent Similarity:	45.45%	Conservative:	22
Best Local Similarity:	27.27%	Mismatches:	49
Query Match:	10.63%	Indels:	17
DB:	3	Gaps:	5

US-09-757-415A-1_COPY_11_140 (1-130) x US-09-103-840A-1 (1-4411529)

Qy	16	ValAspAspGlyAsnGluLeuGlySerGlyIleMetGluLeuThrAspThrGluLeu	35
Db	2085929	GTGACACTGACGCGCAGCGAGTGTTCGAAGCATCATGGATGCCCTATTTCACCGGCGTG	2085870
Qy	36	Ile---LeuTyrThrArgLysArgAspSerValLysTrpHisTyrLeuCysLeuArgArg	54
Db	2085869	ATGCCCATCCACGGGCTAAAGCCAGCGACGTCAACGGCGCGTGATC-----	2085822
Qy	55	TyrGlyTyrAspSerAsnLeuPheSerPheGluSerGlyArgArgCysGlnThrGlyGln	74
Db	2085821	-----AACACGCCGACCGGCTCCATCTACATCGTCAAGCCGAAGATGCACGGTCCG	2085771
Qy	75	GlyIlePheAlaPheLysCysAlaArgAlaGluGluLeuPheAsnMetLeuGlnGluIle	94
Db	2085770	GCCGAGGTGGCGTTTACCTGC-----GAACTGTTCAGCCGGGTTGAAGATGTG	2085723
Qy	95	Met-----GlnAsnAsn---SerIleAsnValValGluGluPro-ValValGluAr	110
Db	2085722	CTGGGGTTGCCGCAAAACACCATGAAGATCGGCATCATGACAGGAGNACCCCGGACCACG	2085663
Qy	110	gAsnAsnHisGlnThrGluLeuGluValProArgThrProArgThrProThrThrProGl	130
Db	2085662	GTCAACCTCAAGGCGGTGCATCAAAAGCTGCCGCGGACCGCGTGTGTTCATCAACACCGGG	2085603
Qy	130	Y 130	
Db	2085602	T 2085602	

Search completed: August 9, 2004, 00:57:52
Job time : 2545.21 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 8, 2004, 23:07:35 ; Search time 582.434 Seconds
(without alignments)
1094.387 Million cell updates/sec

Title: US-09-757-415A-1_COPY_11_140

Perfect score: 696
Sequence: 1 DTVPNHRNKKFKVINDDG.....NNHQTELEVPRTPTPTPG 130

Scoring table:

BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 3222919 seqs, 2451570024 residues

Total number of hits satisfying chosen parameters: 6445838

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US09757415/runat_02082004_101234_6802/app_query.fasta_1.526
-DB=Published Applications NA -QFMT=fastap -SUFFIX=p2n.rnpb -MINMATCH=0.1
-LOOPCL=0 -LOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS-human40.cdi -LISP=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR WFN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09757415 -CGN_1_1_630=runat_02082004_101234_6802
-NCPU=6 -ICPU=3 -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSFBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications NA:
1: /cgn2_6/prodata/1/pubpna/US07_PUBCOMB.seq:
2: /cgn2_6/prodata/1/pubpna/PCT_NEW_PUB.seq:
3: /cgn2_6/prodata/1/pubpna/US06_NEW_PUB.seq:
4: /cgn2_6/prodata/1/pubpna/US06_PUBCOMB.seq:
5: /cgn2_6/prodata/1/pubpna/US07_NEW_PUB.seq:
6: /cgn2_6/prodata/1/pubpna/PCTUS_PUBCOMB.seq:
7: /cgn2_6/prodata/1/pubpna/US08_NEW_PUB.seq:
8: /cgn2_6/prodata/1/pubpna/US08_PUBCOMB.seq:
9: /cgn2_6/prodata/1/pubpna/US09A_PUBCOMB.seq:
10: /cgn2_6/prodata/1/pubpna/US09B_PUBCOMB.seq:
11: /cgn2_6/prodata/1/pubpna/US09C_PUBCOMB.seq:
12: /cgn2_6/prodata/1/pubpna/US09_NEW_PUB.seq:
13: /cgn2_6/prodata/1/pubpna/US09_NEW_PUB.seq:
14: /cgn2_6/prodata/1/pubpna/US10A_PUBCOMB.seq:
15: /cgn2_6/prodata/1/pubpna/US10B_PUBCOMB.seq:
16: /cgn2_6/prodata/1/pubpna/US10C_PUBCOMB.seq:
17: /cgn2_6/prodata/1/pubpna/US10_NEW_PUB.seq:
18: /cgn2_6/prodata/1/pubpna/US60_NEW_PUB.seq:
19: /cgn2_6/prodata/1/pubpna/US60_PUBCOMB.seq:
20: /cgn2_6/prodata/1/pubpna/US60_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
------------	-------	-------	--------	-------	-------------

1	696	100.0	1532	15	US-10-146-473-27	Sequence 27, Appl
2	696	100.0	2074	13	US-10-276-774-842	Sequence 842, App
3	264.5	38.0	452	10	US-09-918-995-27390	Sequence 27390, A
4	143.5	20.6	1908	13	US-10-342-887-601	Sequence 601, App
5	143.5	20.6	1908	13	US-10-172-118-601	Sequence 601, App
6	143.5	20.6	1972	13	US-10-240-425-380	Sequence 380, App
7	143.5	20.6	3809	14	US-10-001-870-68	Sequence 68, Appl
8	141	20.3	1539	9	US-09-789-919-15	Sequence 15, Appl
9	141	20.3	1650	9	US-09-789-919-57	Sequence 57, Appl
10	137.5	19.8	1146	15	US-10-037-270-854	Sequence 854, App
11	137.5	19.8	1146	16	US-10-117-722-854	Sequence 854, App
12	137.5	19.8	1750	16	US-10-307-928A-5	Sequence 5, Appl
13	135.5	19.5	1221	16	US-10-350-923B-37	Sequence 37, Appl
14	133.5	19.2	914	15	US-10-037-270-855	Sequence 855, App
15	133.5	19.2	914	16	US-10-117-722-855	Sequence 855, App
16	124.5	17.9	1951	14	US-10-098-841-102	Sequence 102, App
17	124.5	17.9	2000	10	US-09-814-353-21341	Sequence 21341, A
18	124.5	17.9	2735	15	US-10-037-270-842	Sequence 842, App
19	124.5	17.9	2735	16	US-10-117-722-842	Sequence 842, App
20	122	17.5	421	10	US-09-918-995-34825	Sequence 34825, A
21	120.5	17.3	2339	13	US-10-112-944-41	Sequence 41, Appl
22	113	16.2	508	10	US-09-918-995-24418	Sequence 24418, A
23	111.5	16.0	1385	15	US-10-106-698-1983	Sequence 1983, Ap
24	95.5	13.7	714	15	US-10-029-386-24084	Sequence 24084, A
25	95.5	13.7	804	15	US-10-029-386-24133	Sequence 24133, A
26	95.5	13.7	1409	16	US-10-364-049-165	Sequence 165, App
27	88	12.6	507	15	US-10-029-386-10377	Sequence 10377, A
28	85	12.2	24740	13	US-10-382-844-41	Sequence 41, Appl
29	85	12.2	24740	15	US-10-292-951-41	Sequence 91, Appl
30	83	11.9	1813	15	US-10-114-170-91	Sequence 6, Appl
31	78	11.2	5828	9	US-09-903-248-6	Sequence 6, Appl
32	78	11.2	5828	9	US-09-859-604-6	Sequence 6, Appl
33	78	11.2	5828	9	US-09-903-063-6	Sequence 6, Appl
34	78	11.2	5828	9	US-09-903-216-6	Sequence 6, Appl
35	78	11.2	5828	9	US-09-903-199-6	Sequence 6, Appl
36	78	11.2	5828	9	US-09-880-107-3021	Sequence 3021, Ap
37	78	11.2	5828	9	US-09-903-023-6	Sequence 6, Appl
38	78	11.2	5828	10	US-09-436-184-6	Sequence 6, Appl
39	78	11.2	5828	13	US-10-342-887-1165	Sequence 1165, Ap
40	78	11.2	5828	13	US-10-172-118-1165	Sequence 1165, Ap
41	78	11.2	5828	17	US-10-641-643-1386	Sequence 1386, Ap
42	78	11.2	6321	16	US-10-334-143-155	Sequence 155, App
43	77.5	11.1	698	9	US-09-835-992A-44	Sequence 44, Appl
44	76.5	11.0	95960	13	US-10-087-192-1384	Sequence 1384, Ap
45	76.5	11.0	95960	13	US-10-087-192-1390	Sequence 1390, Ap

ALIGNMENTS

RESULT 1

US-10-146-473-27
Sequence 27, Application US/10146473
Publication No. US2003010888A1
GENERAL INFORMATION:
APPLICANT: Scanlan, Matthew
APPLICANT: Gout, Ivan
APPLICANT: Stockert, Elisabeth
APPLICANT: Gure, Ali
APPLICANT: Chen, Yao-Tseng
APPLICANT: Old, Lloyd
TITLE OF INVENTION: Breast Cancer Antigens
FILE REFERENCE: L00461/70130(JRV)
CURRENT APPLICATION NUMBER: US/10/146,473
CURRENT FILING DATE: 2002-05-15
PRIOR APPLICATION NUMBER: US 60/291,150
PRIOR FILING DATE: 2001-05-15
NUMBER OF SEQ ID NOS: 82
SOFTWARE: PatentIn version 3.0
SEQ ID NO 27
LENGTH: 1532
TYPE: DNA
ORGANISM: Homo sapiens


```
Db 162 ATCAATGTGATGAGAGCCTGTTCATCATCACCCGCAATAGCACCCCGCTGAGCTTGAC 221
Qy 119 ValProArgThrProArgThrProThrThrProGly 130
Db 222 CTCCTCGAGCCGCCAGCACCCCAATGCTCTAGGC 257

RESULT 4
US-10-342-887-601
; Sequence 601, Application US/10342887
; Publication No. US20040038340A1
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linsley, Peter S.
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Christopher J.
; APPLICANT: Van 't Veer, Laura Johanna
; APPLICANT: Van de Vijver, Marc J.
; APPLICANT: Bernards, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-188-999
; CURRENT APPLICATION NUMBER: US/10/342,887
; PRIOR FILING DATE: 2003-01-15
; PRIOR APPLICATION NUMBER: 60/298,918
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: 60/380,710
; PRIOR FILING DATE: 2002-05-14
; PRIOR APPLICATION NUMBER: 10/172,118
; PRIOR FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 601
; LENGTH: 1908
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-342-887-601

Alignment Scores:
Pred. No.: 3,77e-10 Length: 1908
Score: 143.50 Matches: 33
Percent Similarity: 47.32% Conservative: 20
Best Local Similarity: 29.46% Mismatches: 50
Query Match: 20.62% Indels: 9
DB: 13 Gaps: 2

US-09-757-415A-1_COPY_11_140 (1-130) x US-10-342-887-601 (1-1908)
Qy 24 GlySerGlyIleMetGluLeuThrAspThrGluLeuIleLeuTyThrArgLysArgAsp 43
Db 536 GGCTCTACGTGCTGAGGGTGGAGGCTGAAGGCTGACTCTCTGACCGTGGGGGCCAG 595
Qy 44 Ser-----ValLysTrpHisTyLeuCysLeuArgArgTyGlyTyr 57
Db 596 AGTCAGATACCTGGAGCCACTCTCTCTGGCCCTACACTCTGTTCGCTCGCTATGCGCGG 655
Qy 58 AspSerAsnLeuPheSerPheGluSerGlyArgCysGlnThrGlyGlnGlyIlePhe 77
Db 656 GACAAAGGTCATCTTCTCTTCGAGCGCGCGCCCTGCCCTCAGGCCCTGGAACCTTC 715
Qy 78 AlaPheLysCysAlaArgAlaGluLeuPheAsnMetLeuGlnGluIleMetGlnAsn 97
Db 716 ACCTTCAGAGCGGCACAGGGAATGACATCTTCAGGCAGTTGAGACTGCCATCCACCGG 775
Qy 98 AsnSerIleAsnValValGlu-----GluProValValGluArgAsnHisGln 114
Db 776 CAGAAAGCCCGGAGAAAGCGGACAGCGGACGAGTGTTCAGAGCTGACTCCCATGAA 835
Qy 115 ThrGluLeuGluValProArgThrProArgThrPro 126
Db 836 GGGAGGTGGCAGAGGGGAAGTTCCTTCCCACT 871

RESULT 5
US-10-342-887-601
; Sequence 601, Application US/10342887
; Publication No. US20040038340A1
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linsley, Peter S.
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Christopher J.
; APPLICANT: Van 't Veer, Laura Johanna
; APPLICANT: Van de Vijver, Marc J.
; APPLICANT: Bernards, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-188-999
; CURRENT APPLICATION NUMBER: US/10/342,887
; PRIOR FILING DATE: 2003-01-15
; PRIOR APPLICATION NUMBER: 60/298,918
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: 60/380,710
; PRIOR FILING DATE: 2002-05-14
; PRIOR APPLICATION NUMBER: 10/172,118
; PRIOR FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 601
; LENGTH: 1908
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-342-887-601

Alignment Scores:
Pred. No.: 3,77e-10 Length: 1908
Score: 143.50 Matches: 33
Percent Similarity: 47.32% Conservative: 20
Best Local Similarity: 29.46% Mismatches: 50
Query Match: 20.62% Indels: 9
DB: 13 Gaps: 2

US-09-757-415A-1_COPY_11_140 (1-130) x US-10-342-887-601 (1-1908)
Qy 24 GlySerGlyIleMetGluLeuThrAspThrGluLeuIleLeuTyThrArgLysArgAsp 43
Db 536 GGCTCTACGTGCTGAGGGTGGAGGCTGAAGGCTGACTCTCTGACCGTGGGGGCCAG 595
Qy 44 Ser-----ValLysTrpHisTyLeuCysLeuArgArgTyGlyTyr 57
Db 596 AGTCAGATACCTGGAGCCACTCTCTCTGGCCCTACACTCTGTTCGCTCGCTATGCGCGG 655
Qy 58 AspSerAsnLeuPheSerPheGluSerGlyArgCysGlnThrGlyGlnGlyIlePhe 77
Db 656 GACAAAGGTCATCTTCTCTTCGAGCGCGCGCCCTGCCCTCAGGCCCTGGAACCTTC 715
Qy 78 AlaPheLysCysAlaArgAlaGluLeuPheAsnMetLeuGlnGluIleMetGlnAsn 97
Db 716 ACCTTCAGAGCGGCACAGGGAATGACATCTTCAGGCAGTTGAGACTGCCATCCACCGG 775
Qy 98 AsnSerIleAsnValValGlu-----GluProValValGluArgAsnHisGln 114
Db 776 CAGAAAGCCCGGAGAAAGCGGACAGCGGACGAGTGTTCAGAGCTGACTCCCATGAA 835
Qy 115 ThrGluLeuGluValProArgThrProArgThrPro 126
Db 836 GGGAGGTGGCAGAGGGGAAGTTCCTTCCCACT 871

RESULT 6
US-10-240-425-380
; Sequence 380, Application US/10240425
; Publication No. US20040033502A1
; GENERAL INFORMATION:
; APPLICANT: Williams, Amanda
; APPLICANT: Boland, Joseph F.
; APPLICANT: Lord, Reginald V.
; APPLICANT: Alvarez, Chris
; APPLICANT: Wetzel, Jon C.
; APPLICANT: Scherf, Uwe
; APPLICANT: Vockley, Joseph G.
; TITLE OF INVENTION: Gene Expression Profiles in Esophageal Tissue
```

```
; Publication No. US20030224374A1
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linsley, Peter
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Chris
; APPLICANT: Van 't Veer, Laura
; APPLICANT: Van de Vijver, Marc
; APPLICANT: Bernards, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-175-999
; CURRENT APPLICATION NUMBER: US/10/172,118
; CURRENT FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/380,770
; PRIOR FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 601
; LENGTH: 1908
; TYPE: DNA
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: NM_001381
; DATABASE ENTRY DATE: 2001-06-18
US-10-172-118-601

Alignment Scores:
Pred. No.: 3,77e-10 Length: 1908
Score: 143.50 Matches: 33
Percent Similarity: 47.32% Conservative: 20
Best Local Similarity: 29.46% Mismatches: 50
Query Match: 20.62% Indels: 9
DB: 13 Gaps: 2

US-09-757-415A-1_COPY_11_140 (1-130) x US-10-172-118-601 (1-1908)
Qy 24 GlySerGlyIleMetGluLeuThrAspThrGluLeuIleLeuTyThrArgLysArgAsp 43
Db 536 GGCTCTACGTGCTGAGGGTGGAGGCTGAAGGCTGACTCTCTGACCGTGGGGGCCAG 595
Qy 44 Ser-----ValLysTrpHisTyLeuCysLeuArgArgTyGlyTyr 57
Db 596 AGTCAGATACCTGGAGCCACTCTCTCTGGCCCTACACTCTGTTCGCTCGCTATGCGCGG 655
Qy 58 AspSerAsnLeuPheSerPheGluSerGlyArgCysGlnThrGlyGlnGlyIlePhe 77
Db 656 GACAAAGGTCATCTTCTCTTCGAGCGCGCGCCCTGCCCTCAGGCCCTGGAACCTTC 715
Qy 78 AlaPheLysCysAlaArgAlaGluLeuPheAsnMetLeuGlnGluIleMetGlnAsn 97
Db 716 ACCTTCAGAGCGGCACAGGGAATGACATCTTCAGGCAGTTGAGACTGCCATCCACCGG 775
Qy 98 AsnSerIleAsnValValGlu-----GluProValValGluArgAsnHisGln 114
Db 776 CAGAAAGCCCGGAGAAAGCGGACAGCGGACGAGTGTTCAGAGCTGACTCCCATGAA 835
Qy 115 ThrGluLeuGluValProArgThrProArgThrPro 126
Db 836 GGGAGGTGGCAGAGGGGAAGTTCCTTCCCACT 871

RESULT 6
US-10-240-425-380
; Sequence 380, Application US/10240425
; Publication No. US20040033502A1
; GENERAL INFORMATION:
; APPLICANT: Williams, Amanda
; APPLICANT: Boland, Joseph F.
; APPLICANT: Lord, Reginald V.
; APPLICANT: Alvarez, Chris
; APPLICANT: Wetzel, Jon C.
; APPLICANT: Scherf, Uwe
; APPLICANT: Vockley, Joseph G.
; TITLE OF INVENTION: Gene Expression Profiles in Esophageal Tissue
```

```
FILE REFERENCE: 44921-5026
; CURRENT APPLICATION NUMBER: US/10/240,425
; CURRENT FILING DATE: 2002-09-30
; PRIOR APPLICATION NUMBER: PCT/US01/09847
; PRIOR FILING DATE: 2001-03-28
; PRIOR APPLICATION NUMBER: US 60/193,446
; PRIOR FILING DATE: 2000-03-31
; NUMBER OF SEQ ID NOS: 1588
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 380
; LENGTH: 1972
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20040033502A1 AF035299
US-10-240-425-380

Alignment Scores:
Pred. No.: 3,97e-10 Length: 1972
Score: 143.50 Matches: 33
Percent Similarity: 47.32% Conservative: 20
Best Local Similarity: 29.46% Mismatches: 50
Query Match: 20.62% Indels: 9
DB: 13 Gaps: 2

US-09-757-415A-1_COPY_11_140 (1-130) x US-10-240-425-380 (1-1972)
QY 24 GlySerGlyIleMetGluLeuThrAspThrGluLeuIleLeuTyrThrArgLysArgAsp 43
Db 600 GGCTCTACGTCGTGAGGGTGAGGCTGAAAGGCTGACTCTCTCGACCGCGGGGGCCCG 659
QY 44 Ser-----ValLysTrpHisTyrLeuCysLeuArgArgTyrGlyTyr 57
Db 660 AGTCAGATACCTGAGCCACTCTCTCTCGCCCTACACATCTGTGTGCGTCGCTATGCGCG 719
QY 58 AspSerAsnLeuPheSerGlyArgArgCysGlnThrGlyGlnGlyIlePhe 77
Db 720 GACAGGTCATGTTCTTTTCAGCGCGCGCGCGCTGCGCCCTCAGCGCCCTGGAACCTTC 779
QY 78 AlaPheLysCysAlaArgAlaGluLeuPheAsnMetLeuGlnGluIleMetGlnAsn 97
Db 780 ACCTTCCAGCGGCACAGGGAATGACATCTTCCAGGCAGTTGAGCTGCATCCACCGG 839
QY 98 AsnSerIleAsnValValGlu-----GluProValValGluArgAsnAsnHisGln 114
Db 840 CAGAAGGCCCGAAGGCGGACAGGCGGACGATGTTCTCAGAGCTGACTCCCAATGAA 899
QY 115 ThrGluLeuGluValProArgThrProArgThrPro 126
Db 900 GGGGAGGTGGCAGAGGGAAGTTGCTTCCCTCCCACT 935

RESULT 7
US-10-001-870-68
; Sequence 68, Application US/10001870
; Publication No. US20020150924A1
; GENERAL INFORMATION:
; APPLICANT: Salceda, Susana
; APPLICANT: Macina, Roberto
; APPLICANT: Recipon, Herve
; APPLICANT: Sun, Yongming
; APPLICANT: Liu, Chenghua
; TITLE OF INVENTION: Compositions and Methods Relating to Prostate Specific Genes and
; FILE REFERENCE: DEX-0283
; CURRENT APPLICATION NUMBER: US/10/001,870
; CURRENT FILING DATE: 2001-11-20
; PRIOR APPLICATION NUMBER: 60/252,189
; PRIOR FILING DATE: 2000-11-21
; NUMBER OF SEQ ID NOS: 217
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 68
; LENGTH: 3809
; TYPE: DNA
; ORGANISM: Homo sapien

US-09-757-415A-1_COPY_11_140 (1-130) x US-10-001-870-68 (1-3809)
QY 24 GlySerGlyIleMetGluLeuThrAspThrGluLeuIleLeuTyrThrArgLysArgAsp 43
Db 2397 GGCTCTACGTCGTGAGGGTGAGGCTGAAAGGCTGACTCTCTCGACCGTGGGGCCCG 2456
QY 44 Ser-----ValLysTrpHisTyrLeuCysLeuArgArgTyrGlyTyr 57
Db 2457 AGTCAGATACCTGAGCCACTCTCTCTCGCCCTACACATCTGTGTGCGTCGCTATGCGCG 2516
QY 58 AspSerAsnLeuPheSerGlyArgArgCysGlnThrGlyGlnGlyIlePhe 77
Db 2517 GACAAAGTCATGTTCTTTCAGCGCGCGCGCTGCGCCCTCAGGCGCTGGAACCTTC 2576
QY 78 AlaPheLysCysAlaArgAlaGluLeuPheAsnMetLeuGlnGluIleMetGlnAsn 97
Db 2577 ACCTTCCAGCGGCACAGGGAATGACATCTTCCAGGCAGTTGAGCTGCATCCACCGG 2636
QY 98 AsnSerIleAsnValValGlu-----GluProValValGluArgAsnAsnHisGln 114
Db 2637 CAGAAGGCCCGAAGGCGGACAGGCGGACGATGTTCTCAGAGCTGACTCCCAATGAA 2696
QY 115 ThrGluLeuGluValProArgThrProArgThrPro 126
Db 2697 GGGGAGGTGGCAGAGGGAAGTTGCTTCCCTCCCACT 2732

RESULT 8
US-09-789-919-15
; Sequence 15, Application US/09789919
; Patent No. US2002006485A1
; GENERAL INFORMATION:
; APPLICANT: Lemischka, Ihor
; APPLICANT: Moore, Kateri
; TITLE OF INVENTION: GENES THAT REGULATE HEMATOPOIETIC BLOOD FORMING STEM
; FILE REFERENCE: 2275-1-005
; CURRENT APPLICATION NUMBER: US/09/789,919
; CURRENT FILING DATE: 2001-02-21
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 15
; LENGTH: 1539
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-789-919-15

Alignment Scores:
Pred. No.: 6,33e-10 Length: 1539
Score: 141.00 Matches: 36
Percent Similarity: 44.09% Conservative: 20
Best Local Similarity: 28.35% Mismatches: 47
Query Match: 20.26% Indels: 24
DB: 9 Gaps: 3

US-09-757-415A-1_COPY_11_140 (1-130) x US-09-789-919-15 (1-1539)
QY 23 LeuGlySerGlyIleMetGluLeuThrAspThrGluLeuIleLeuTyrThrArgLysArg 42
Db 560 CTGGCCCAAGATGACATCCAACTGAGGAGACA-----TCCAGCCCGG 604
QY 43 AspSerValLysTrpHisTyrLeuCysLeuArgArgTyrGlyTyrAspSerAsnLeuPhe 62
Db 605 GCCTCTTTTAGTGGCCCTACCGTTCTCGCGCAAGTACGCTCTGACAGGGGTGTGTT 664
```

Qy 63 SerPheGluSerGlyArgArgCysGlnThrGlyGlnGlyLeuPheAlaPheLysCysAla 82
Db 665 TCGTTTGGAGCTGGCCGCTGTGACTCAGGTGAGGGCTTTTTCCTTCAGTAGCCCG 724
Qy 83 ArgAlaGluGluPheAsnMetLeuGlnGluLeuMetGlnAsnSerIleAsnVal 102
Db 725 CGTGCCCCAGACATATGTGGGTGTGGCTGCCGCCATTGCCCGCCAGGAGCGTCTT 784
Qy 103 ValGlu----- 104
Db 785 CCAGAGCTGGCCATGTCCCCACCGCTGCCCTGCTCGGCCCTCTCCCTGCCCTCCCTA 844
Qy 105 GluProValValGluArgAsnAsnHisGlnThrGluLeuGluValProArgThrProArg 124
Db 845 GAGCCCTCTGAGAGCTTCGGGAGGTGGCCCGCAGGATTGAGCTGCCACTCCACAGAG 904
Qy 125 ThrPro---ThrThrProGly 130
Db 905 CTGCTCTTAAGTATGCCCGG 925
RESULT 9
US-09-789-919-57
; Sequence 57, Application US/09789919
; Patent No. US20020064855A1
; GENERAL INFORMATION:
; APPLICANT: Lemischka, Ibor
; APPLICANT: Moore, Kateri
; TITLE OF INVENTION: GENES THAT REGULATE HEMATOPOIETIC BLOOD FORMING STEM
; FILE OF INVENTION: CELLS AND USES THEREOF
; CURRENT APPLICATION NUMBER: US/09789,919
; CURRENT FILING DATE: 2001-02-21
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 57
; LENGTH: 1650
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-789-919-57
Alignment Scores:
Pred. No.: 7,05e-10 Length: 1650
Score: 141.00 Matches: 36
Percent Similarity: 44.09% Conservative: 20
Best Local Similarity: 28.35% Mismatches: 47
Query Match: 20.26% Indels: 24
DB: 9 Gaps: 3
US-09-757-415A-1_COPY_11_140 (1-130) x US-09-789-919-57 (1-1650)
Qy 23 LeuGlySerGlyIleMetGluLeuThrAspThrGluLeuLeuLeuLeuThrArgLysArg 42
Db 671 CTGGCCCAAGATGACATCCCACTCAGGAGACA-----TCCAGCCCGCAG 715
Qy 43 AspSerValLysTrpHisTyrLeuCysLeuArgArgTyrGlyTyrAspSerAsnLeuPhe 62
Db 716 GCCTGTTTACCTGCGCTACCGTTTCTGCGCAAGTACCGCTCTGCAAGGGGTGTCTC 775
Qy 63 SerPheGluSerGlyArgArgCysGlnThrGlyGlnGlyLeuPheAlaPheLysCysAla 82
Db 776 TCGTTTGGAGCTGGCCGCTGTGACTCAGGTGAGGGCTTTTTCCTTCAGTAGCCCG 835
Qy 83 ArgAlaGluGluPheAsnMetLeuGlnGluLeuMetGlnAsnSerIleAsnVal 102
Db 836 CGTGCCCCAGACATATGTGGGTGTGGCTGCCGCCATTGCCCGCCAGGAGCGTCTT 895
Qy 103 ValGlu----- 104
Db 896 CCAGAGCTGGCCATGTCCCCACCGCTGCCCTGCTCGGCCCTCTCCCTGCCCTCCCTA 955
Qy 105 GluProValValGluArgAsnAsnHisGlnThrGluLeuGluValProArgThrProArg 124

Db 956 GAGCCCCCTGGAGAGCTTCGGGAGTGGCCCCCAGGATTGTAGCTGCCCACTCCACAGAAAG 1015
Qy 125 ThrPro---ThrThrProGly 130
Db 1016 CTGCTCTTAAGTATGCCCGG 1036
RESULT 10
US-10-037-270-854
; Sequence 854, Application US/10037270
; Publication No. US20030104529A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yungqing
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: Tillinghast, John
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. US20030104529A1el Nucleic Acids and
; FILE OF INVENTION: Polypeptides
; FILE REFERENCE: 784CIP23
; CURRENT APPLICATION NUMBER: US/10/037,270
; CURRENT FILING DATE: 2002-01-04
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1104
; SOFTWARE: pt_FL_Genes Version 1.0
; SEQ ID NO 854
; LENGTH: 1146
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (109)..(1122)
US-10-037-270-854
Alignment Scores:
Pred. No.: 1.32e-09 Length: 1146
Score: 137.50 Matches: 34
Percent Similarity: 49.07% Conservative: 19
Best Local Similarity: 31.48% Mismatches: 42
Query Match: 19.76% Indels: 13
DB: 15 Gaps: 3
US-09-757-415A-1_COPY_11_140 (1-130) x US-10-037-270-854 (1-1146)
Qy 24 GlySerGlyIleMetGluLeuThrAspThrGluLeuLeuLeuLeuThrArgLysArgAsp 43
Db 565 GGTGAATGCACAAATGCAGATCACTCATGAAATAATCTATCTCTGGGATATCCCAATGCC 624
Qy 44 SerValLys-----TrpHisTyrLeuCysLeuArgArgTyrGlyTyrAspSerAsn 60
Db 625 AGGTCAAACTGGTGTATGTGGCTCTCAGCTCAGTACAGGAGATAGGTGGGACTCAACG 684
Qy 61 LeuPheSerPheGluSerGlyArgArgCysGlnThrGlyGlnGlyLeuPheAlaPheLys 80
Db 685 TGGTTTCACTTTGAGTCAGGAAGAATGTGTGACACAGGAGAGGACTATTCTACTTTTCAA 744
Qy 81 CysAlaArgAlaGluLeuPheAsnMetLeuGlnGluLeuMetGlnAsnSerIle 100
Db 745 ACAAGGGAAGGAGAA-----ATGATCTATCAGAGGTTTCATTCTCGGACACTG 792


```
Db 653 GGTGAATGCACAAATGCAGATCACTCATGAAATAATATCTATCTCTGGGATATCCACAAATGCC 712
Qy 44 SerValLys-----TrrHisTyrLeuCysLeuArgTyrGlyTyrAspSerAsn 60
Db 713 AGGTCAAACTGGTGTATGGCTCTCAGCTCAGTGGAGATACGGTCGGACTCAACG 772
Qy 61 LeuPheSerPheGluSerGlyArgArgCysGlnThrGlyGlnGlyLeuPheAlaPheLys 80
Db 773 TGGTTTCAGCTTGGAGTCAGGAAGATGTGTGACACAGAGGAGGACTATTCTACTTTTCAA 832
Qy 81 CysAlaArgAlaGluLeuPheAsnMetLeuGlnGluLeuMetClnAsnAsnSerIle 100
Db 833 ACAAGGAAGGAGAA-----ATGATCTATCAGAAGGTTCAITCTCGCACACTG 880
Qy 101 AsnValValGluGlu-----ProValValGluArgAsnAsnHisGln 114
Db 881 GCCATAGCTGAGCAACATGAAGATTATGCTAGAAATGGAACAGAGAGCCCGGCTTCAG 940
Qy 115 ThrGluLeuValProArgThr 122
Db 941 ACAAGCTTGACTGAACCAATGACA 964

RESULT 13
US-10-350-923B-37
; Sequence 37, Application US/10350923B
; Publication No. US20040002084A1
; GENERAL INFORMATION:
; APPLICANT: Dotson, Stanton B.
; TITLE OF INVENTION: Nucleic acid, polypeptides, vectors, and cells derived from activ
; FILE REFERENCE: S03157-01
; CURRENT APPLICATION NUMBER: US/10/350,923B
; PRIOR FILING DATE: 2003-01-24
; PRIOR APPLICATION NUMBER: US/ 09/454,280
; PRIOR FILING DATE: 1999-12-03
; PRIOR APPLICATION NUMBER: PCT/US99/28773
; PRIOR FILING DATE: 1999-06-12
; PRIOR APPLICATION NUMBER: US 60/111,006
; PRIOR FILING DATE: 1998-12-04
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 37
; LENGTH: 1221
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-350-923B-37

Alignment Scores:
Pred. No.: 2,87e-09 Length: 1221
Score: 135.50 Matches: 37
Percent Similarity: 49.07% Conservative: 16
Best Local Similarity: 34.26% Mismatches: 35
Query Match: 19.47% Indels: 20
DB: 16 Gaps: 4

US-09-757-415A-1_COPY_11_140 (1-130) x US-10-350-923B-37 (1-1221)
Qy 38 TyrThrArgLysArgAsp----SerValLys-TrrHisTyrLeuCysLeuArgTyrGln 56
Db 348 TGGCCCGGAGCCAGGACCCAGCTGTACGACTGTCAGGTTCTCTGGCGGCTTTGG 407
Qy 56 YTyAspSerAsnLeuPheSerPheGluSerGlyArgArgCysGlnThrGlyGlnGly 76
Db 408 GCGGGAACAAGTAACCTTTTCCTTTGAGCGAGGCGCTCGTGGTCTCTCGAGAGGCGAA 467
Qy 76 ePheAlaPheLysCysAlaArgAlaGluGluLeuPheAsnMetClnGluLeuMet-- 95
Db 468 CTITGAGTTCGAACCCGGAAGCAATGAGATCTTCTTGGCCCTTGGAGAGCCCATCTC 527
Qy 96 ----GlnAsnAsnSerIleAsnValValGluGluProValValGluArgAsnHisGln 114
```

```
Db 528 TGCCAGAGAAGTCTGTCACCCGCTACACCCCAACCG-----CAGCCAGGCAC 575
Qy 114 nThrGluLeuGluValProArg-----Th 122
Db 576 AATCCCGCGCTGCTGCCCGGCTGATAGCCCTACTCTCTGGCGCATGACTCACTGCC 635
Qy 122 rProArgThrProThrPro 129
Db 636 GCCGCCTTACCCCAACACCG 657

RESULT 14
US-10-037-270-855
; Sequence 855, Application US/10037270
; Publication No. US20030104529A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yungqing
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: Tillinghaast, John
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. US20030104529A1el Nucleic Acids and
; FILE REFERENCE: 784CIP2B
; CURRENT APPLICATION NUMBER: US/10/037,270
; CURRENT FILING DATE: 2002-01-04
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1104
; SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 855
; LENGTH: 914
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (109)..(867)
US-10-037-270-855

Alignment Scores:
Pred. No.: 3,62e-09 Length: 914
Score: 133.50 Matches: 33
Percent Similarity: 48.62% Conservative: 20
Best Local Similarity: 30.28% Mismatches: 43
Query Match: 19.18% Indels: 13
DB: 15 Gaps: 3

US-09-757-415A-1_COPY_11_140 (1-130) x US-10-037-270-855 (1-914)
Qy 24 GlySerGlyIleMetGluLeuThrAspThrGluLeuIleLeuTyrThrArgLysArgAsp 43
Db 565 GGTGAATGCACAAATGCAGATCACTCATGAAATAATATCTCTGGGATATCCACAAATGCC 624
Qy 44 SerValLys-----TrrHisTyrLeuCysLeuArgTyrGlyTyrAspSerAsn 60
Db 625 AAGTCAAACTGGTGTATGGCTCTCAGCTCAGTCACTGAGAGATACGGTCGGACTCAACG 684
Qy 61 LeuPheSerPheGluSerGlyArgArgCysGlnThrGlyGlnGlyLeuPheAlaPheLys 80
Db 685 TGGTTTCAGCTTGGAGTCAGGAAGATGTGTGACACAGAGGAGGACTATTCTACTTTTCAA 744
```

QY 81 CysAlaArgAlaGluGluLeuPheAsnMetLeuGlnGluMetGlnAsnAsnSerIle 100
|||
Db 745 ACAAGGAAGAGAA-----ATGATCTATCAGAAGGTTTCATTCTGCACACTG 792
QY 101 AsnValValGluGluProValValGluArgAsnAsnHisGlnThrGluLeuGluValPro 120
|||
Db 793 GCCATAGCTGAGCAA-----CATGAAAGATTAAATGCTAGAAATGGAA 834
QY 121 ArgThrProArgThrProThrThrPro 129
|||
Db 835 CAGAAGGCCCGGGTAAGGCCCTTCCT 861

RESULT 15

US-10-117-722-855
; Sequence 855, Application US/10117722
; Publication No. US20030219744A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. US20030219744A1el Nucleic Acids and
; FILE REFERENCE: 784CIP2BCIP
; CURRENT APPLICATION NUMBER: US/10/117,722
; CURRENT FILING DATE: 2002-04-04
; PRIOR APPLICATION NUMBER: 09/620,312
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1104
; SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 855
; LENGTH: 914
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (109)..(867)
US-10-117-722-855

Alignment Scores:
Pred. No.: 3,62e-09 Length: 914
Score: 133.50 Matches: 33
Percent Similarity: 48.62% Conservative: 20
Best Local Similarity: 30.28% Mismatches: 43
Query Match: 19.18% Indels: 13
DB: 16 Gaps: 3

US-09-757-415a-1_COPY_11_140 (1-130) x US-10-117-722-855 (1-914)

QY 24 GlySerGlyIleMetGluLeuThrAspThrGluLeuIleLeuThrArgLysArgAsp 43
|||
Db 565 GGTGAATGCACAAATGCAGATCACTCATGAAAATATCTCTCTGGGATATCCCAATGCC 624
QY 44 SerValLys-----TTPHisTyrLeuCysLeuArgArgTyrGlyTyrAspSerAsn 60
|||
Db 625 AAGTCAAACTGGTGATGTGGCTCTCAGTCTACTGAGGAGATACCGTGGGACTCAACG 584
QY 61 LeuPheSerPheGluSerGlyArgArgCysGlnThrGlyGlnGlyIlePheAlaPheLys 80
|||
Db 685 TGGTTACAGTTTCAGTCAGGAAGAATGTGTACACACAGGAGAGACTATTCTCTTCAA 744
QY 81 CysAlaArgAlaGluGluLeuPheAsnMetLeuGlnGluIleMetGlnAsnAsnSerIle 100
|||
Db 745 ACAAGGAAGAGAA-----ATGATCTATCAGAAGGTTTCATTCTGCACACTG 792
QY 101 AsnValValGluGluProValValGluArgAsnAsnHisGlnThrGluLeuGluValPro 120
|||

Db 793 GCCATAGCTGAGCAA-----CATGAAAGATTAAATGCTAGAAATGGAA 834

QY 121 ArgThrProArgThrProThrThrPro 129
|||
Db 835 CAGAAGGCCCGGGTAAGGCCCTTCCT 861

Search completed: August 9, 2004, 02:11:10
Job time : 591.434 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 2, 2004, 09:22:38 ; Search time 13.8947 Seconds
(without alignments)
447.367 Million cell updates/sec

Title: US-09-757-415A-3

Perfect score: 107

Sequence: 1 HSQMAVHKLAKSLPRRQVTVS 22

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A Geneseq_29Jan04.*

1: Geneseqp1980s.*

2: Geneseqp1980s.*

3: Geneseqp2000s.*

4: Geneseqp2001s.*

5: Geneseqp2002s.*

6: Geneseqp2003as.*

7: Geneseqp2003bs.*

8: Geneseqp2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	107	100.0	22	4 AAU04695	AAU04695 Mouse fib
2	107	100.0	702	5 ABG79680	ABG79680 Tumour in
3	107	100.0	731	2 AAR10649	AAR10649 Chicken b
4	107	100.0	731	2 AAR21686	AAR21686 Human bFG
5	107	100.0	764	3 AAB58376	AAB58376 Lung canc
6	107	100.0	816	2 AAR39647	AAR39647 Human fib
7	107	100.0	820	2 AAR47233	AAR47233 Human fib
8	107	100.0	820	2 AAU63844	AAU63844 Mouse bFG
9	107	100.0	820	2 AAU06458	AAU06458 Fibroblas
10	107	100.0	820	4 AAB84383	AAB84383 Amino aci
11	107	100.0	822	2 AAR13549	AAR13549 Basic FGF
12	107	100.0	822	2 AAR20750	AAR20750 flg recep
13	107	100.0	822	2 AAR26337	AAR26337 N-sam. 2/
14	107	100.0	822	4 AAU04694	AAU04694 Mouse fib
15	107	100.0	822	6 AAE34488	AAE34488 Human FGF
16	107	100.0	822	7 ABR56164	ABR56164 Human Fib
17	107	100.0	822	7 ADE61698	ADE61698 Rat Prote
18	107	100.0	822	7 ADE55803	ADE55803 Rat Prote
19	107	100.0	822	7 ADE61700	ADE61700 Human Pro
20	97	90.7	819	2 AAR10648	AAR10648 Chicken b
21	83	77.6	643	2 AAR15267	AAR15267 Clone pTB
22	83	77.6	654	7 ADD47011	ADD47011 Human Pro
23	83	77.6	726	2 AAR14280	AAR14280 Murine KG
24	83	77.6	726	2 AAR10933	AAR10933 KGF recep
25	83	77.6	769	2 AAR15268	AAR15268 Clone pTB

26	83	77.6	821	2 AAR21080	AAR21080 flg recep
27	83	77.6	821	7 AD557113	AD557113 Human Pro
28	83	77.6	821	7 ADD45446	ADD45446 Human Pro
29	81	75.7	729	2 AAW37804	AAW37804 Fibroblas
30	81	75.7	729	2 AAY23631	AAY23631 Bek-like
31	81	75.7	731	2 AAY23633	AAY23633 Newt Kera
32	81	75.7	750	2 AAW37806	AAW37806 Keratinoc
33	71	66.4	526	2 AAR15266	AAR15266 Clone pTB
34	71	66.4	526	2 AAR15269	AAR15269 Clone pTB
35	71	66.4	679	2 AAR41517	AAR41517 K-sam. 3/
36	64	59.8	20	6 ABR98369	ABR98369 Tumour ce
37	64	59.8	20	7 ADC84715	ADC84715 MCF-7 bre
38	50	46.7	122	4 ABG29386	ABG29386 Novel hum
39	47	43.9	4391	6 AAE34390	AAE34390 Human per
40	47	43.9	4393	4 AAB31889	AAB31889 Amino aci
41	47	43.9	4436	4 ABG23265	ABG23265 Novel hum
42	46	43.0	577	4 ABG23262	ABG23262 Novel hum
43	44	41.1	194	4 AAG89977	AAG89977 C glutami
44	44	41.1	1520	7 ADC86177	ADC86177 Human GPC
45	43	40.2	277	6 ABU23813	ABU23813 Protein e

ALIGNMENTS

RESULT 1

AAU04695
ID AAU04695 standard; peptide; 22 AA.

XX AAU04695;

AC AC

DT 26-SEP-2001 (first entry)

XX Mouse fibroblast growth factor receptor-1 (FGFR1)-derived peptide.

XX Suc-1-associated neurotrophic factor target protein; SNT-1;

KW protein co-ordinate data; mouse; fibroblast growth factor receptor; FGFR;

KW nuclear magnetic resonance; NMR spectroscopy; rational drug design;

XX three-dimensional structure.

XX Mus musculus.

XX WO200151521-A2.

XX 19-JUL-2001.

PD 10-JAN-2001; 2001WO-US000821.

XX 12-JAN-2000; 2000US-0175867P.

PR 09-JAN-2001; 2001US-00757415.

XX (MOUN) MOUNT SINAI SCHOOL MEDICINE.

XX Zhou M;

PI WPI; 2001-425868/45.

XX New isolated nucleic acid sequences encoding polypeptides useful in

DR structure based rational drug design.

XX Claim 11; Page 44; 235pp; English.

XX The sequence represents the amino acid sequence of mouse fibroblast

CC growth factor receptor-1 (FGFR1)-derived peptide. The invention provides

CC fragments of suc-1-associated neurotrophic factor target protein (SNT)

CC and fibroblast growth factor receptor (FGFR) which can form a binding

CC complex that is amenable to structural determinations by nuclear magnetic

CC resonance (NMR) spectroscopy. The invention provides methodology for

CC related structure based rational drug design using the three-dimensional

CC data. The invention fulfills the need for the determination of the three-

CC dimensional structure of such complexes as described and for procedures

CC for related structural base drug design predicated on such structural

CC data

XX SQ Sequence 22 AA;
 Query Match 100.0%; Score 107; DB 4; Length 22;
 Best Local Similarity 100.0%; Pred. No. 1.9e-11;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HSQMAVHKLAKSIPLRRQVTVS 22
 |||||
 Db 1 HSQMAVHKLAKSIPLRRQVTVS 22

RESULT 2
 ABG79680
 ID ABG79680 standard; protein; 702 AA.
 AC ABG79680;
 DT 15-NOV-2002 (first entry)
 DE Tumour involved gene (TIG) splice variant protein, NV-11.
 XX Human; splice variant; tumour-involved gene; TIG;
 KW pharmaceutical composition; cancer; diagnostic; tumour; gene therapy;
 KW endothelial cell; cell differentiation; cell proliferation; apoptosis;
 KW Gene therapy.
 XX Homo sapiens.
 OS US2002086384-A1.
 PN 04-JUL-2002.
 PD 13-MAR-2001; 2001US-00805020.
 XX 14-MAR-2000; 2000IL-00135402.
 PR 16-MAY-2000; 2000IL-00136154.
 XX (LEVI/) LEVINE Z.
 PA (DAVI/) DAVID A.
 PA (ROMA/) ROMANO C.
 PA (BERN/) BERNSTEIN J.
 XX Levine Z, David A, Romano C, Bernstein J;
 WPI: 2002-635679/68.
 DR N-PSDB; ABS65210.
 XX Novel nucleic acid sequence, which is an alternative splicing variant of
 PT tumor involved genes, useful for detecting cancer, predisposition to
 PT cancer, for evaluating cancer state and in gene therapy for treating
 PT cancer.
 XX Claim 4; Page 76-77; 180pp; English.

XX The invention discloses isolated human nucleic acid alternative splicing
 CC variants that are all tumour-involved genes (TIGs). The nucleic acids and
 CC polypeptides are useful for determining the level of a nucleic acid or
 CC polypeptide in a biological sample, for detecting a variant nucleic acid
 CC or polypeptide sequence in a biological sample, for determining the level
 CC of variant nucleic acid or polypeptide sequences in a biological sample
 CC and for determining the ratio between the level of variant sequence from which
 CC first biological sample and the level of the original sequence from which
 CC the variant has been varied by alternative splicing in a second
 CC biological sample and for raising antibodies. A pharmaceutical
 CC composition comprising a carrier and the nucleic acid, is useful for
 CC treating diseases (e.g. cancer) that can be ameliorated or cured by
 CC increasing or decreasing the level of the encoded protein. The nucleic
 CC acids are also useful for diagnostic purposes, especially for detecting
 CC cancer or a predisposition to cancer, for evaluating the state or
 CC aggressiveness of cancer disease, in basic research, for understanding
 CC the physiological function of the original TIG, in targeting or
 CC developing pharmaceuticals, for distinguishing various stages in the life

XX cycle of the same type of cells which may be helpful for the development
 CC of pharmaceuticals for various cancer stages in which cell cycle is non-
 CC normal, for determining mutations in tumour-involved genes and in gene
 CC therapy. The polypeptides are useful for identifying compounds capable of
 CC binding to the variant product and modulating its activity and for
 CC modulating endothelial differentiation and proliferation, as well as to
 CC modulate apoptosis either ex vivo or in vivo. The sequences presented in
 CC ABG796700-ABG79705 are the new variants (NV) 1-36 proteins of the TIGs
 CC disclosed
 XX SQ Sequence 702 AA;
 Query Match 100.0%; Score 107; DB 5; Length 702;
 Best Local Similarity 100.0%; Pred. No. 9.9e-10;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HSQMAVHKLAKSIPLRRQVTVS 22
 |||||
 Db 409 HSQMAVHKLAKSIPLRRQVTVS 430

RESULT 3
 AAR10649
 ID AAR10649 standard; protein; 731 AA.
 XX AAR10649;
 AC AAR10649;
 XX 24-OCT-2003 (revised)
 DT 25-MAR-2003 (revised)
 DT 17-APR-1991 (first entry)
 XX Chicken basic fibroblast growth factor.
 DE FGF; angiogenesis; neuron growth.
 KW Gallus gallus.
 OS WO9100916-A.
 PN 24-JAN-1991.
 PD 06-JUL-1989; 89US-00377003.
 PF 06-JUL-1989; 89US-00377003.
 PR (REGC) UNIV CALIFORNIA.
 XX Williams LT, Johnson DE, Lee PE;
 PI WPI; 1991-051340/07.
 DR P-PSDB; AAR10648.
 XX Fibroblast growth factor receptors - and clones encoding them, sol. forms
 PT of receptors and antibodies to them useful for diagnostic studies and
 PT therapy.
 XX Disclosure; Fig 3; 84pp; English.

XX The sequence was deduced from a cDNA clone isolated from a library prepd.
 CC from mRNA from chicken embryos (day 6). It encodes a basic fibroblast
 CC growth factor receptor which can be used for tissue repair or
 CC development and for neuron maintenance. The DNA and protein and
 CC antibodies prepd. from it can be used to treat diseases resulting from
 CC aberrant FGF expression and for elucidating the mechanism of FGF and
 CC identifying agonists for the prevention of angiogenesis and control of
 CC tumours. See also AAR10648. (Updated on 25-MAR-2003 to correct PA field.)
 CC (Updated on 24-OCT-2003 to standardise OS field)
 XX SQ Sequence 731 AA;
 Query Match 100.0%; Score 107; DB 2; Length 731;
 Best Local Similarity 100.0%; Pred. No. 1e-09;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HSQMAVHKLAKSIPLRRQVTVS 22
 DB 318 HSQMAVHKLAKSIPLRRQVTVS 339

RESULT 4
 AAR21686
 ID AAR21686 standard; protein; 731 AA.

XX AC AAR21686;
 XX DT 03-JUL-1992 (first entry)
 XX DE Human bFGF receptor protein.
 XX KW Basic fibroblast growth factor; fms like gene.
 XX OS Homo sapiens.
 XX PN JP04036184-A.

XX PD 06-FEB-1992.

XX PF 01-JUN-1990; 90JP-00141490.

XX PR 01-JUN-1990; 90JP-00141490.

XX PA (MITK) MITSUI TOATSU CHEM INC.

XX DR WPI; 1992-092893/12.

XX DR N-PSDB; AAR20914.

XX PT Human b-FGF receptor gene - contains recombinant DNA, chromosome DNA or
 XX PS cDNA base sequence encoding polypeptide.
 XX Claim 1; Fig 2; 10pp; Japanese.

XX The human basic fibroblast growth factor (b-FGF) receptor protein
 CC sequence was deduced from the DNA sequence obtd. from a human placenta
 CC cDNA library in lambda gill. The library was screened with two probes
 CC (AAQ24634,5) which were derived from the known FLG sequence (fms like
 CC gene). Positive clones were transformed into E. coli Y1090 strain, and
 CC recombinant phase were sequenced

XX Sequence 731 AA;

Query Match 100.0%; Score 107; DB 2; Length 731;
 Best Local Similarity 100.0%; Pred. No. 1e-09;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HSQMAVHKLAKSIPLRRQVTVS 22
 DB 318 HSQMAVHKLAKSIPLRRQVTVS 339

RESULT 5
 AAB58376
 ID AAB58376 standard; protein; 764 AA.

XX AC AAB58376;

XX DT 14-MAR-2001 (first entry)

XX DE Lung cancer associated polypeptide sequence SEQ ID 714.

XX Human; lung cancer associated protein; neuroprotective; cytostatic;
 KW cardiactive; immunomodulatory; muscular active; vulnerary;
 KW gastrointestinal; nephrotropic; antinefactive; gynecological;
 KW antibacterial; diagnosis; neural disorder; immune disorder; reproductive;
 KW proliferative disorder; wound healing; infectious disease.

XX OS Homo sapiens.

XX PN WO200055180-A2.
 XX DT 21-SEP-2000.
 XX PF 08-MAR-2000; 2000WO-US0005918.
 XX PR 12-MAR-1999; 99US-0124270P.
 XX PA (HUMA-) HUMAN GENOME SCI INC.
 XX PA (ROSE/) ROSEN C A.
 XX PI Ruben SM;
 XX WPI; 2000-587514/55.
 XX DR N-PSDB; AAF18252.
 XX PT Lung cancer associated gene sequences, referred to as lung cancer
 PT antigens, useful for treatment, prevention, and diagnosis of disorders
 PT such as lung cancer.
 XX Claim 11; Page 1225-1228; 1425pp; English.

XX Polynucleotide sequences AAF17982 - AAF18424 encode human lung cancer
 CC associated proteins represented in AAB58106 - AAB58548. Lung cancer
 CC associated proteins and polynucleotide sequences, their agonists, and
 CC antagonists may have neuroprotective; cytostatic; cardioactive;
 CC immunomodulatory; muscular active general; vulnerary; gastrointestinal
 CC general; nephrotropic; antinefactive; gynecological; or antibacterial
 CC activity. The invention also includes antibodies specific for the protein
 CC or polynucleotide sequences. The lung cancer associated polynucleotide
 CC sequences may be used for detection of lung cancer, chromosome
 CC identification, as chromosome markers, and for numerous other diagnostic
 CC or research purposes. The proteins may be used to treat disorders such as
 CC neural, immune, muscular, reproductive, gastrointestinal, pulmonary,
 CC cardiovascular, renal, and proliferative disorders. The proteins may also
 CC be used in the treatment of wounds and infectious diseases.

XX Polynucleotide sequences AAF18425 - AAF18433 and peptide AAB58549 are
 CC used in the course of the invention for the identification and
 CC characterisation of the polynucleotide and protein sequences

XX Sequence 764 AA;

Query Match 100.0%; Score 107; DB 3; Length 764;
 Best Local Similarity 100.0%; Pred. No. 1.1e-03;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HSQMAVHKLAKSIPLRRQVTVS 22

DB 351 HSQMAVHKLAKSIPLRRQVTVS 372

RESULT 6
 AAR39647
 ID AAR39647 standard; protein; 816 AA.

XX AC AAR39647;

XX DT 14-JAN-1994 (first entry)

XX DE Human fibroblast growth factor receptor.

XX soluble; FGF-R; fibroblast growth factor receptor; tumour therapy;
 KW immunoglobulin-like domain; tyrosine kinase activity; hyperplasia;
 KW psoriasis; Herpes Simplex Virus; HSV.

XX OS Homo sapiens.

XX Key Location/Qualifiers

XX FT Cleavage-site 21..22

XX FT Domain /note= "putative signal peptidase site"

XX FT 55..101

XX FT /label= Ig-like_domain_1

```

XX CS 126. .133
XX FT /label= ARR
XX FT /note= "acidic amino acid rich region"
FH FT Misc-difference 137. .143
FT FT EEKE in Fig. 2"
FT FT 179. .231
FT FT /label= Ig-like_domain-2
FT FT 278. .341
FT FT /label= Ig-like_domain-3
FT FT 372. .392
FT FT /label= transmembrane
FT FT 475. .575
FT FT /label= Tyrosine Kinase domain
FT FT /note= "first part of split TK domain"
FT FT 590. .750
FT FT /label= Tyrosine Kinase domain
FT FT /note= "second part of split TK domain"
XX FT
XX PN US5229501-A.
XX PD 20-JUL-1993.
XX PF 11-JAN-1991; 91US-00640029.
XX PR 11-JAN-1991; 91US-00640029.
XX PA (CHIR ) CHIRON CORP.
XX PI Keifer MC, Valenzuela PDT, Barr PU;
XX PS WPI; 1993-242535/30.
XX CC The invention covers both the full-length FGF-receptor and its soluble
XX CC form (i.e. truncated by deletion of the C-terminal portion beginning at
XX CC the transmembrane domain). The FGF receptor is defined as having the
XX CC sequence in Columns 13-18 and Fig 2. These sequences differ slightly as
XX CC indicated by the "Misc difference" in the Features Table. The receptor
XX CC can be used to detect FGF or to inhibit FGF activity. Specifically, the
XX CC soluble form of the receptor can inhibit FGF-dependent tumour growth,
XX CC angiogenesis, psoriasis, excessive scar formation and hyperplasias and
XX CC can be used to treat Herpes Simplex virus infections
XX SQ Sequence 816 AA;
XX CC Query Match 100.0%; Score 107; DB 2; Length 816;
XX CC Best Local Similarity 100.0%; Pred. No. 1.2e-09; Indels 0; Gaps 0;
XX CC Matches 2; Conservative 0; Mismatches 0;
QY 1 HSQMAVHKLAKSIPLRRQVTVS 22
DB 404 HSQMAVHKLAKSIPLRRQVTVS 425
RESULT 7
AAR47233
ID AAR47233 standard; protein; 820 AA.
XX AC AAR47233;
XX DT 25-MAR-2003 (revised)
XX DT 07-SEP-1994 (first entry)
XX DE Human fibroblast growth factor receptor.
XX KW Cytomegalovirus; CMV; Towne; GH gene; escort protein;
XX KW recombinant protein production; viral glycoprotein H; FGF receptor;
XX KW fibroblast growth factor receptor; soluble; flg5 cDNA clone.

```

```

XX CS Homo sapiens.
XX FT Key
FH FT Peptide
FT FT /label= signal_peptide
FT FT 22. .820
FT FT /label= human_FGF_receptor
FT FT 55. .101
FT FT /label= immunoglobulin-like_domain_1
FT FT /note= "extracellular"
FT FT 126. .133
FT FT /label= ARR
FT FT /note= "acidic amino acid-rich region"
FT FT 176. .228
FT FT /label= immunoglobulin-like_domain_2
FT FT /note= "extracellular"
FT FT 275. .339
FT FT /label= immunoglobulin-like_domain_3
FT FT /note= "extracellular"
FT FT 374. .395
FT FT /label= transmembrane_domain
FT FT 478. .578
FT FT /label= TK
FT FT /note= "first part of split intracellular tyrosine kinase
FT FT domain"
FT FT 593. .754
FT FT /label= TK
FT FT /note= "second part of split intracellular tyrosine
FT FT kinase domain"
XX PN WO9403620-A2.
XX PD 17-FEB-1994.
XX PF 29-JUL-1993; 93WO-US007299.
XX PR 29-JUL-1992; 92US-00921807.
XX CC (CHIR ) CHIRON CORP.
XX CC Spaete RL;
XX CC WPI; 1994-065708/08.
XX CC Increased expression and secretion of viral protein e.g. from
XX CC cytomegalovirus from host cells - by co-expression with DNA encoding an
XX CC escort protein, e.g. fibroblast growth factor receptor.
XX CC Claim 8; Fig 6; 86pp; English.
XX CC Expression of recombinant viral glycoproteins on host cell surfaces can
XX CC be significantly enhanced by coexpression with an escort protein. The CMV
XX CC glycoprotein H polypeptide (see AAR47232) is a preferred viral
XX CC glycoprotein and is pref. coexpressed with human fibroblast growth factor
XX CC receptor. A full-length FGF receptor cDNA, designated flg5, was isolated
XX CC from a lambda ZAP human hepatoma HepG2 cDNA library; flg5 codes for a 820
XX CC amino acid protein having the sequence AAR47233. (Updated on 25-MAR-2003
XX CC to correct PN field.)
XX SQ Sequence 820 AA;
XX CC Query Match 100.0%; Score 107; DB 2; Length 820;
XX CC Best Local Similarity 100.0%; Pred. No. 1.2e-09; Indels 0; Gaps 0;
XX CC Matches 22; Conservative 0; Mismatches 0;
QY 1 HSQMAVHKLAKSIPLRRQVTVS 22
DB 407 HSQMAVHKLAKSIPLRRQVTVS 428
RESULT 8
AAW63844

```

AAW63844 standard; protein; 820 AA.
 AAW63844;
 25-MAR-2003 (revised)
 01-OCT-1998 (first entry)
 Mouse bFGF receptor protein.
 Basic fibroblast growth factor receptor; bFGF; heparin binding; murine;
 antitumour agent; inhibitor; wound healing.
 Mus sp.
 Key Location/Qualifiers
 Misc-difference 67 /label= unknown
 Misc-difference 68 /label= unknown
 Misc-difference 260 /label= unknown
 Misc-difference 281 /label= unknown
 Misc-difference 282 /label= unknown
 Misc-difference 576 /label= unknown
 Misc-difference 702 /label= unknown
 US5789182-A.
 04-AUG-1998.
 14-DEC-1993; 93US-00166717.
 20-DEC-1990; 90US-00631717.
 (HARD) HARVARD COLLEGE.
 (CHIL-) CHILDRENS MEDICAL CENT.
 Leder P, Yayon A, Flanagan JG, Ornitz DM, Klagsbrun M;
 WPI; 1998-446074/38.
 N-PSDB; AAV44041.
 Assays for high-affinity heparin-binding growth factor receptor ligands -
 using receptor-overexpressing cells or cell-free system.
 Disclosure; Fig 9; 38pp; English.
 This sequence represents a murine basic fibroblast growth factor (bFGF)
 which is a member of the heparin-binding growth factor receptor family.
 This protein is used in a method which assays the ability of a substance
 to bind to a high-affinity heparin-binding growth factor (HBGF) receptor.
 The assay screens for potential antitumour agents that inhibit binding of
 HBGF to high-affinity receptors, or for potential wound healing agents
 that promote such binding. (Updated on 25-MAR-2003 to correct PF field.)
 Sequence 820 AA;
 Query Match 100.0%; Score 107; DB 2; Length 820;
 Best Local Similarity 100.0%; Pred. No. 1.2e-09;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 HSQMAVHKLAKSIPLRRQVTWS 22
 DB 407 HSQMAVHKLAKSIPLRRQVTWS 428
 RESULT 9
 AAY06458
 ID AAY06458 standard; protein; 820 AA.

XX AAY06458;
 AC 27-SEP-1999 (first entry)
 DT Fibroblast growth factor receptor 1.
 XX ZNF198-FGFR1; fibroblast growth factor receptor 1; oncogene; lymphoma;
 KW leukaemia; SCLL; translocation; stem cell; diagnosis; therapy.
 XX Homo sapiens.
 XX Key Location/Qualifiers
 FH Misc-difference 426..427
 FT /note= "translocation breakpoint"
 FT Domain 483..488
 FT /note= "tyrosine kinase domain I"
 XX WO9935159-A1.
 XX 15-JUL-1999.
 XX 07-JAN-1999; 99WO-US000365.
 XX 08-JAN-1999; 98US-00004688.
 XX (BGHM) BRIGHAM & WOMENS HOSPITAL.
 XX Fletcher JA, Xiao S;
 PI WPI; 1999-430381/36.
 XX N-PSDB; AAX87179.
 XX Oncogene, designated ZNF198-FGFR1, useful in diagnosis of stem cell
 leukemia and lymphoma syndrome.
 PS Disclosure; Page 70-72; 76pp; English.
 XX The present sequence represents human fibroblast growth factor receptor 1
 (FGFR1), which is encoded by a gene on human chromosome 8. A newly
 identified oncogene, ZNF198-FGFR1 (see AAX87177), incorporates a FGFR1
 tyrosine kinase domain fused through t(8;13) translocation to ZNF198, a
 novel zinc finger protein (see AAY06457) and is associated with stem cell
 leukaemia/lymphoma syndrome (SCLL). Molecular characterisation of the
 oncogene and the encoded oncoprotein (see AAY06456) provides the ability
 to identify patients with t(8;13)(p11;q11-12) syndrome prior to onset of
 myeloid leukaemia. The oncogene can be identified with primers and probes
 capable of hybridizing to ZNF198 or FGFR1 tyrosine kinase domain coding
 sequences. Cells containing ZNF198-FGFR1 can be located using a
 radiolabeled antibody that recognizes the fusion junction locus.
 CC Antibodies that recognize this fusion junction can also be used to
 deliver toxic substances to patients containing the ZNF198-FGFR1 protein.
 CC This is useful for treatment of SCLL
 XX Sequence 820 AA;
 Query Match 100.0%; Score 107; DB 2; Length 820;
 Best Local Similarity 100.0%; Pred. No. 1.2e-09;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 HSQMAVHKLAKSIPLRRQVTWS 22
 DB 407 HSQMAVHKLAKSIPLRRQVTWS 428
 RESULT 10
 AAB84383
 ID AAB84383 standard; protein; 820 AA.
 XX AAB84383;
 AC AAB84383;
 XX 22-AUG-2001 (first entry)
 DT
 XX

DE Amino acid sequence of a fibroblast growth factor receptor.
 KW Human; fibroblast growth factor receptor; FGFR, wound healing;
 malignancy.

OS Homo sapiens.

PN US6255454-B1.

PD 03-JUL-2001.

XX 12-MAY-1995; 95US-00439992.

PR 11-JAN-1991; 91US-00640029.

PR 12-APR-1993; 93US-00046020.

PR 30-SEP-1994; 94US-00315686.

PA (CHIR) CHIRON CORP.

PI Keifer MC, Valenzuela PDT, Barr PJ;

DR WPI; 2001-388930/41.

XX Novel fragment of a human fibroblast growth factor receptor, used to

PT detect fibroblast growth factor in samples and to identify modulators of

PT the factor, which are used to accelerate wound healing.

XX Claim 1; Col 13-18; 20pp; English.

XX The present sequence represents a human fibroblast growth factor receptor

CC (FGFR). FGFR is useful for detecting FGF in a sample, and for inhibiting

CC the activity of FGF. The FGFR may also be used to screen and identify

CC compounds which mimic FGF binding. FGFR can be used to identify agonists

CC and antagonists. The agonists identified may be used to accelerate wound

CC healing, and the antagonists identified can be used to inhibit the growth

CC of FGF-related malignancies

XX Sequence 820 AA;

Query Match 100.0%; Score 107; DB 4; Length 820;

Best Local Similarity 100.0%; Pred. No. 1.2e-09;

Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HSQMAVHKLAKSIPLRRQVTVS 22

DB 407 HSQMAVHKLAKSIPLRRQVTVS 428

RESULT 11

AAR13549

ID AAR13549 standard; protein; 822 AA.

XX AAR13549;

XX 25-MAR-2003 (revised)

DT 28-OCT-1991 (first entry)

XX Basic FGF receptor.

PD 08-AUG-1991.

XX 23-JAN-1990; 90GB-00001466.

XX 23-JAN-1990; 90GB-00001466.

XX (FARM) FARMITALIA ERBA SRL CARLO.

XX Bergonzoni L, Mazue G, Isacchi A, Roncucci R, Sarmientos P;

XX WPI; 1991-252611/34.

DR N-PSDB; AAQ13311.

XX Extracellular form of human fibroblast growth factor receptor - used to

PT treat tumours, abnormal angiogenesis e.g. diabetic retinopathy,

PT rheumatoid arthritis and arteriosclerosis and as contraceptives.

XX Claim 1; Fig 3; 29pp; English.

XX The sequence was deduced from the cDNA sequence obtd. from two

CC overlapping clones, PLS and PLS10, isolated from a placental lambda gt11

CC cDNA library. The DNA can be used to express recombinant bFGF receptor

CC which is an antagonist of human aFGF and bFGF. The receptor can be used

CC to treat abnormal angiogenesis (e.g. in diabetic retinopathy, neovascular

CC glaucoma, etc, and possibly certain solid tumours), and in

CC contraceptives. Dosage is 10-100 ug. (Updated on 25-MAR-2003 to correct

CC PA field.)

XX Sequence 822 AA;

Query Match 100.0%; Score 107; DB 2; Length 822;

Best Local Similarity 100.0%; Pred. No. 1.2e-09;

Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HSQMAVHKLAKSIPLRRQVTVS 22

DB 409 HSQMAVHKLAKSIPLRRQVTVS 430

RESULT 12

AAR20750

ID AAR20750 standard; protein; 822 AA.

XX AAR20750;

XX 25-MAR-2003 (revised)

DT 20-MAY-1992 (first entry)

XX flg receptor protein.

XX Fibroblast growth factor receptor; heparin binding proteins;

XX tyrosine kinase; fms-like gene; CSF-1; PDGF.

XX Homo sapiens.

XX Key Location/Qualifiers

FT Peptide 1..21

FT Protein /label= signal_peptide

FT Modified-site 22..822

FT 185 /label= flg_receptor_protein

FT /note= "N-linked glycosylation site"

XX WO9200999-A.

XX 23-JAN-1992.

XX 06-JUL-1990; 90US-00549587.

XX 06-JUL-1990; 90US-00549587.

XX (RORE) RORER INT HOLDINGS INC.

PI Dionne CA, Crumley G, Jaye MC, Schlessing J;
XX WPI; 1992-056827/07.
DR N-PSDB; AAQ21003.
XX
XX New fibroblast growth factor receptor proteins - useful in treating gf-
PT mediated conditions e.g. angiogenesis of tumours, mitogenic effects in
PT psoriasis, arthritis.
XX
XX Claim 5; Fig 7; 65pp; English.
XX
XX The protein sequence was deduced from the DNA sequence of an flg cDNA
CC clone having increased binding for aFGF receptors. Flg and bek are
CC similar but distinct gene products with structural features shared by the
CC pDGF/CSF-1/c-kit family of receptor linked tyrosine kinases. They have a
CC 21 residue hydrophobic signal sequence, flg has an extracellular domain
CC of 355 residues and a cytoplasmic domain of 425 residues. The
CC extracellular domain contains 3 "immunoglobulin like" domains of similar
CC size and location. There are nine potential N-glycosylation sites. There
CC is an "acidic box" of 8 amino acids. The cytoplasmic domain consists of
CC long juxtamembrane regions followed by catalytic kinase domains which are
CC split by 14 amino acid insertions. The kinase domains are followed by
CC divergent carboxy terminal tails. Receptor proteins encoded by this cDNA
CC sequence may be used in pharm-aceutical compen. to inhibit undesirable
CC heparin-binding growth factor mediated cellular responses or to inhibit
CC the binding of an opportunistic pathogen to human cells. Such undesirable
CC responses may be growth factor stimulated angiogenesis and
CC vascularisation of tumours, mitogenic effects in psoriasis, arthritis,
CC atherosclerosis and benign prostatic hyper- trophy. The derived flg
CC protein may also be used for screening drugs for treatment of such
CC problems. See also AAR21080 (Updated on 25-MAR-2003 to correct PA field.)
XX
XX Sequence 822 AA;
SQ
Query Match 100.0%; Score 107; DB 2; Length 822;
Best Local Similarity 100.0%; Pred. No. 1.2e-09;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 HSQMAVHKLAKSIPLRRQVTVS 22
DB 409 HSQMAVHKLAKSIPLRRQVTVS 430
RESULT 13
AAR26337
ID AAR26337 standard; protein; 822 AA.
XX
XX AAR26337;
AC
XX 03-FEB-1993 (first entry)
DT
XX N-sam.
DE
XX Tumour gene; bFGF receptor; antibody; detection; drug.
XX Homo sapiens.
XX JP04190732-A.
XX 09-JUL-1992.
XX
XX 22-NOV-1990; 90JP-00316100.
XX 22-NOV-1990; 90JP-00316100.
XX (TAKE) TAKEDA CHEM IND LTD.
XX (KOKU-) KOKURITSU GAN CENT SOCHO.
XX WPI; 1992-280111/34.
XX N-PSDB; AAQ27658.
XX
XX Gene prod of N-sam tumour - and its recombinant-contg. nucleotide
PT sequence.

XX Claim 1; Fig 1; 19pp; Japanese.
XX
XX The sequence given encoded by the N-sam tumour gene. This protein could
CC be the receptor of human bFGF. The N-sam protein and peptide fragments of
CC N-sam are useful as anti-tumour drugs. Anti-N-sam peptide antibodies are
CC useful for the detection of the N-sam gene product, in the purification
CC of the N-sam gene product and in the diagnosis of tumours and anti-tumour
CC drugs
XX
XX Sequence 822 AA;
SQ
Query Match 100.0%; Score 107; DB 2; Length 822;
Best Local Similarity 100.0%; Pred. No. 1.2e-09;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 HSQMAVHKLAKSIPLRRQVTVS 22
DB 409 HSQMAVHKLAKSIPLRRQVTVS 430
RESULT 14
AAU04694
ID AAU04694 standard; protein; 822 AA.
XX
XX AAU04694;
AC
XX 26-SEP-2001 (first entry)
DT
XX Mouse fibroblast growth factor receptor-1 (FGFR1).
XX
XX Suc-1-associated neurotrophic factor target protein; SNT-1;
KW protein co-ordinate data; mouse; fibroblast growth factor receptor; FGFR;
KW nuclear magnetic resonance; NMR spectroscopy; rational drug design;
KW three-dimensional structure.
XX
XX Mus musculus.
OS
XX WC200151521-A2.
XX
XX 19-JUL-2001.
PD
XX
XX 10-JAN-2001; 2001WO-US000821.
PF
XX 12-JAN-2000; 2000US-0175867P.
XX
XX 09-JAN-2001; 2001US-00757415.
PR
XX (MOUN) MOUNT SINAI SCHOOL MEDICINE.
PA
XX Zhou M;
PI
XX WPI; 2001-425868/45.
DR
XX New isolated nucleic acid sequences encoding polypeptides useful in
PT structure based rational drug design.
PT
XX Example; Page 230-233; 235pp; English.
PS
XX The sequence represents the amino acid sequence of mouse fibroblast
CC growth factor receptor-1 (FGFR1). The invention provides fragments of suc-
CC -1-associated neurotrophic factor target protein (SNT) and fibroblast
CC growth factor receptor (FGFR) which can form a binding complex that is
CC amenable to structural determinations by nuclear magnetic resonance (NMR)
CC spectroscopy. The invention provides methodology for related structure
CC based rational drug design using the three-dimensional data. The
CC invention fulfills the need for the determination of the three-
CC dimensional structure of such complexes as described and for procedures
CC for related structural base drug design predicated on such structural
CC data
XX
XX Sequence 822 AA;
SQ
Query Match 100.0%; Score 107; DB 4; Length 822;

Best Local Similarity 100.0%; Pred. No. 1.2e-09;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HSQMAVHKLAKSIPLRRQVTVS 22
DB 409 HSQMAVHKLAKSIPLRRQVTVS 430

RESULT 15

AAE34488
ID AAE34488 standard; protein; 822 AA.

XX AAE34488;

XX 14-MAY-2003 (first entry)

DE Human FGF-R1 protein.

XX Human; pyridylpyrimidine derivative; cellular protein kinase; Scrapie;
KW cellular protein phosphatase; cellular signal transduction; prophylaxis;
KW prion infection; chronic wasting disease; CWD; Creutzfeldt-Jacob disease;
KW CJD; transmissible mink encephalopathy; bovine spongiform encephalopathy;
KW TME; BSE; Gerstmann-Strausler-Scheinker syndrome; GSS; Alpers syndrome;
KW fatal familial insomnia; FFI; kuru; neurodegenerative disease; nontropic;
KW Alzheimer's disease; FGF-R1; flg; Flt-1; Flt-2; b-FGFR.

XX Homo sapiens.

XX WO200293164-A2.

XX 21-NOV-2002.

XX 16-MAY-2002; 2002WO-EP005420.

XX 16-MAY-2001; 2001EP-00111858.

XX 29-MAY-2001; 2001US-0293528P.

XX 13-JUL-2001; 2001EP-00117113.

XX 18-JUL-2001; 2001US-0305898P.

XX (AXXI-) AXXIMA PHARM AG.

XX Stein-Gerlach M, Salassidis K, Bacher G, Mueller S;

XX WPI: 2003-120714/11.

XX N-PSDB; AAD52780.

XX New pyridylpyrimidine derivatives useful in the treatment or prevention

XX of infectious disease e.g. Kuru syndrome and Creutzfeldt-Jacob disease

XX (CJD).

XX Disclosure; Page 63-65; 96pp; English.

XX The invention relates to novel pyridylpyrimidine derivatives and methods
XX of detecting prion infections and/or prion disease in an individual or in
XX cells, cell cultures and/or cell lysates. The method involves adding at
XX least one monoclonal or polyclonal antibody, oligonucleotide or pyridyl-
XX pyrimidine derivative to the sample or in cells, cell cultures and/or
XX cell lysates and detecting the activity of at least one human cellular
XX protein kinases (e.g. FGF-R1 (also known as flg, Flt-1, Flt-2, b-FGFR),
XX Tkt (also known as CCK-2, DDR-2 or EDDR; EC number 2.7.1.112), Abl (also
XX known as c-abl), ctk1, MKK7 (also known as SAPK1a, SAPKalpha), CDC2 (also
XX known as CDK1), PRK), human cellular protein phosphatases such as PTP-SL
XX (also known as MCP83) and PTP-zeta, the cellular signal transduction
XX molecules HSP80 and GPR-1. The invention is useful for regulating the
XX production of prions in cells and in the manufacture of pharmaceutical
XX composition for prophylaxis and/or treatment of infectious disease (e.g.
XX Scrapie, chronic wasting disease (CWD), transmissible mink encephalopathy
XX (TME), Creutzfeldt-Jacob disease (CJD), bovine spongiform encephalopathy
XX (BSE), variant CJD, Gerstmann-Strausler-Scheinker syndrome (GSS), fatal
XX familial insomnia (FFI), Kuru and Alpers syndrome, especially BSE, CJD,
XX vCJD) or neurodegenerative diseases (e.g., Alzheimer's disease) in humans
XX or ruminants. The present sequence is human FGF-R1 protein used in the
XX invention

XX SQ Sequence 822 AA;

Query Match 100.0%; Score 107; DB 6; Length 822;
Best Local Similarity 100.0%; Pred. No. 1.2e-09;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HSQMAVHKLAKSIPLRRQVTVS 22

DB 409 HSQMAVHKLAKSIPLRRQVTVS 430

Search completed: August 2, 2004, 09:29:31
Job time : 14.8947 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 2, 2004, 09:26:54 ; Search time 3.32895 Seconds
(without alignments)
635.701 Million cell updates/sec

Title: US-09-757-415A-3
Perfect score: 107
Sequence: 1 HSQMAVHKLAKSIPLRRQVTVS 22

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues
Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 781*
1: Piri: *
2: Piri2: *
3: Piri3: *
4: Piri4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	107	100.0	662	2 C40862	heparin-binding gr
2	107	100.0	733	2 I49293	fibroblast growth
3	107	100.0	822	1 TVHFG	fibroblast growth
4	107	100.0	822	1 TVM3FG	fibroblast growth
5	107	100.0	822	2 S29840	fibroblast growth
6	107	100.0	822	2 I49289	fibroblast growth
7	107	100.0	832	2 JH0393	fibroblast growth
8	97	90.7	814	1 A39752	fibroblast growth
9	97	90.7	819	1 TVCHFG	fibroblast growth
10	95	88.8	729	2 A56795	fibroblast growth
11	94	87.9	812	1 A36477	fibroblast growth
12	83	77.6	415	2 I65223	heparin-binding fi
13	83	77.6	707	2 A54846	fibroblast growth
14	83	77.6	707	2 A38429	xeratinocyte growt
15	83	77.6	769	2 S16236	fibroblast growth
16	83	77.6	820	2 S17295	fibroblast growth
17	83	77.6	821	1 TVMSBK	fibroblast growth
18	83	77.6	821	1 TVHUF2	fibroblast growth
19	83	77.6	822	2 A45081	fibroblast growth
20	83	77.6	822	2 A41794	xeratinocyte growt
21	83	77.6	822	2 B54846	fibroblast growth
22	83	77.6	823	2 B35963	protein-tyrosine k
23	83	77.6	824	2 S24108	protein-tyrosine k
24	82	76.6	824	2 S36439	fibroblast growth
25	81	75.7	748	2 S41050	fibroblast growth
26	81	75.7	750	2 S41051	fibroblast growth
27	73	68.2	816	2 A49151	fibroblast growth
28	71	66.4	682	2 A35969	heparin-binding gr
29	71	66.4	705	2 S51635	fibroblast growth

30	67	62.6	480	2 A56182	fibroblast growth
31	67	62.6	480	2 B56182	fibroblast growth
32	67	62.6	813	1 A49123	fibroblast growth
33	47	43.9	4391	2 A38096	perlecan precursor
34	46	43.0	1829	2 T36135	hypothetical prote
35	45.5	42.5	993	2 S35633	DNA-binding protei
36	45	42.1	3707	2 S18252	heparan sulfate pr
37	44.5	41.6	713	2 I50128	fibroblast growth
38	43.5	40.7	135	1 QQCVC4	AL2 protein - toma
39	43	40.2	217	2 A13429	transposase BME114
40	43	40.2	234	2 AC3205	IS66 family Orf4 (
41	43	40.2	277	2 B97285	ribosomal protein
42	43	40.2	446	2 G32293	phosphoglucutase
43	43	40.2	818	2 JG4058	fibroblast growth
44	42.5	39.7	797	2 S38579	fibroblast growth
45	42	39.3	180	2 H95095	hypothetical prote

ALIGNMENTS

RESULT 1

C40862
heparin-binding growth factor receptor variant alpha-a2 - human
C:Species: Homo sapiens (man)
C:Date: 28-Feb-1992 #sequence_revision 28-Feb-1992 #text_change 16-Jul-1999
C:Accession: C40862
R:Hou, J.; Kan, M.; McKeenhan, K.; McBride, G.; Adams, P.; McKeenhan, W.L.
Science 251, 665-668, 1991
A:Title: Fibroblast growth factor receptors from liver vary in three structural domains.
A:Reference number: A40862; MUID:91126480; PMID:1846977
A:Accession: C40862
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-662 <HOU>
A:Cross-references: GB:M63888
C:Superfamily: basic fibroblast growth factor receptor 1; immunoglobulin homology; protei
C:Keywords: ATP; Growth factor receptor; heparin binding; transmembrane protein
F:171-232/Domain: immunoglobulin homology <IMM>
F:377-399/Domain: transmembrane #status predicted <TMM>
F:476-618/Domain: protein kinase homology #status atypical <KIN>
F:484-492/Region: protein kinase ATP-binding motif

Query Match 100.0%; Score 107; DB 2; Length 662;
Best Local Similarity 100.0%; Pred. No. 1e-09;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HSQMAVHKLAKSIPLRRQVTVS 22
DB 409 HSQMAVHKLAKSIPLRRQVTVS 430

RESULT 2

I49293
fibroblast growth factor receptor-1, short isoform precursor - mouse
C:Species: Mus musculus (house mouse)
C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 16-Jul-1999
C:Accession: I49293
R:Jin, Y.; Pasumamathi, K.B.; Bock, M.E.; Lytras, A.; Kardami, E.; Cattini, P.A.
J. Mol. Cell. Cardiol. 26, 1449-1459, 1994
A:Title: Cloning and expression of fibroblast growth factor receptor-1 isoforms in the m
A:Reference number: I49289; MUID:95205422; PMID:7897669
A:Accession: I49293
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-733 <RES>
A:Cross-references: EMBL:U23445; NID:g733537; PIDN:AAC52183.1; PID:g733538
C:Superfamily: basic fibroblast growth factor receptor 1; immunoglobulin homology; protei
C:Keywords: ATP; Growth factor receptor
F:181-254/Domain: immunoglobulin homology <IMM>
F:387-672/Domain: protein kinase homology <KIN>
F:395-403/Region: protein kinase ATP-binding motif

A;Residues: 81-100 <RUS>
A;Experimental source: recombinant soluble form of extracellular domain after expression
A;Note: this sequence represents the amino end of a 33k fragment protected from trypsin
C;Comment: This receptor binds acidic and basic fibroblast growth factors with high affinity
C;Genetics:
A;Gene: GDB:FGFR1; FLT2
A;Cross-references: GDB:119913; OMIM:136350
A;Map position: 8p11.2-8p11.1
C;Superfamily: basic fibroblast growth factor receptor 1; immunoglobulin homology; prote
C;Keywords: alternative splicing; ATP; autophosphorylation; duplication; glycoprotein; g
yrosine-specific protein kinase
F;1-21/Domain: signal sequence #status predicted <SIG>
F;22-822/Product: fibroblast growth factor receptor 1, form alpha a1 #status predicted <
F;22-427,430-822/Product: fibroblast growth factor receptor 1, form alpha b1 #status pre
F;22-376/Domain: extracellular #status predicted <EXT>
F;22-147,150-822/Product: fibroblast growth factor receptor 1, form 3 #status predicted
F;22-30,120-822/Product: fibroblast growth factor receptor 1, form beta a1 #status predi
F;22-30,120-427,430-822/Product: fibroblast growth factor receptor 1, form beta b1 #stat
F;126-133/Region: acidic
F;161-822/Product: fibroblast growth factor receptor 1, form gamma a1 #status predicted
F;161-427,430-822/Product: fibroblast growth factor receptor 1, form gamma b1 #status pr
F;171-232/Domain: immunoglobulin homology <IMM>
F;377-397/Domain: transmembrane #status predicted <TM>
F;398-822/Domain: intracellular #status predicted <INT>
F;476-761/Domain: protein kinase homology <KIN>
F;484-492/Region: protein kinase ATP-binding motif
F;55-101,178-230,277-341/Bisulfide bonds: #status predicted
F;77,117,227,240,264,296,317,330/Binding site: carboxylate (Asn) (covalent) #status pre
F;514,531,623/Active site: Lys, Glu, Asp #status predicted
F;628,641/Binding site: magnesium (Asn, Asp) #status predicted
F;654/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation) #status predicted

Query Match 100.0%; Score 107; DB 1; Length 822;
Best Local Similarity 100.0%; Pred. No. 1.3e-09;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 HSOQAVHKLAKSIPLRRQVTVS 22
|||||
Db 409 HSOQAVHKLAKSIPLRRQVTVS 430
|||||

RESULT 4
TVM5FG
fibroblast growth factor receptor 1 precursor - mouse
N;Alternate names: basic fibroblast growth factor receptor
N;Contains: protein-tyrosine kinase (SC 2.7.1.112) flg
C;Species: Mus musculus (house mouse)
C;Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 16-Jul-1999
C;Accession: A34849; B34849; S09953; A35794; A43025; PC2277
R;Reid, H.H.; Wilks, A.F.; Bernard, O.
Proc. Natl. Acad. Sci. U.S.A. 87, 1596-1600, 1990
A;Title: Two forms of the basic fibroblast growth factor receptor-like mRNA are expressed
A;Reference number: A34849; MUID:90160373; PMID:1689490
A;Accession: A34849
A;Molecule type: mRNA
A;Residues: 1-822 <REI>
A;Cross-references: GDB:M28998; NID:g192164; PIDN:AAA37290.1; PID:g309116
A;Accession: B34849
A;Molecule type: mRNA
A;Residues: 1-30,120-822 <R2>
A;Cross-references: GDB:M28998
A;Note: the shorter form is expressed at higher levels in neuronal cells at early stages
R;Safarian, A.; Avivi, A.; Orr-Urtreger, A.; Neufeld, G.; Lonai, P.; Givol, D.; Yarden, Y
Oncogene 5, 635-643, 1990
A;Title: The murine flg gene encodes a receptor for fibroblast growth factor.
A;Reference number: S09953; MUID:90265603; PMID:2161096
A;Accession: S09953
A;Molecule type: mRNA
A;Residues: 1-147,150-255, 'ILQ', 259-439, 'A', 441-755, 'R', 757-822 <SAF>
A;Cross-references: EMBL:X51893; NID:g50959; PIDN:CAA36175.1; PID:g50960
R;Masukhani, A.; Moscatelli, D.; Talatlico, D.; Levytska, V.; Basillico, C.
Proc. Natl. Acad. Sci. U.S.A. 87, 4378-4382, 1990
A;Title: A murine fibroblast growth factor (FGF) receptor expressed in CHO cells is acti

A;Reference number: A35794; MUID:90272715; PMID:2161540
A;Accession: A35794
A;Status: nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-30,120-228, 'S', 230-255, 'ILQ', 259-269, 'A', 271-543, 'M', 545-755, 'R', 757-764, 'I'
A;Cross-references: GDB:M33760
R;Basillico, C.
submitted to GenBank, March 1990
A;Reference number: A43025
A;Accession: A43025
A;Molecule type: mRNA
A;Residues: 1-30,120-228, 'S', 230-255, 'ILQ', 259-269, 'A', 271-543, 'M', 545-755, 'R', 757-764, 'I'
A;Cross-references: GDB:M33760; NID:g193298; PIDN:AAA37622.1; PID:g309240
R;Harada, T.; Saito, H.; Kouhara, H.; Kurebayashi, S.; Kasayama, S.; Terakawa, N.; Kishin
Biochem. Biophys. Res. Commun. 205, 1057-1063, 1994
A;Title: Murine fibroblast growth factor receptor 1 gene generates multiple messenger RN
A;Reference number: PC2277; MUID:95100926; PMID:7802632
A;Accession: PC2277
A;Molecule type: DNA
A;Residues: 1-15 <HAR>
A;Cross-references: GDB:S74765; NID:g833887; PIDN:AA332845.1; PID:g833889
C;Comment: this protein mediates the biological actions of heparin-binding growth factors
C;Superfamily: basic fibroblast growth factor receptor 1; immunoglobulin homology; protei
C;Keywords: alternative splicing; ATP; autophosphorylation; duplication; glycoprotein; g
protein kinase
F;1-21/Domain: signal sequence #status predicted <SIG>
F;22-822/Product: fibroblast growth factor receptor 1 #status predicted <MAT>
F;22-376/Domain: extracellular #status predicted <EXT>
F;22-30,120-822/Product: fibroblast growth factor receptor 1, short form #status predicte
F;126-133/Region: acidic
F;171-232/Domain: immunoglobulin homology <IMM>
F;377-397/Domain: transmembrane #status predicted <TM>
F;398-822/Domain: intracellular #status predicted <INT>
F;476-761/Domain: protein kinase homology <KIN>
F;484-492/Region: protein kinase ATP-binding motif
F;55-101,178-230,277-341/Bisulfide bonds: #status predicted
F;77,117,227,240,264,296,317,330/Binding site: carboxylate (Asn) (covalent) #status pre
F;514,531,623/Active site: Lys, Glu, Asp #status predicted
F;628,641/Binding site: magnesium (Asn, Asp) #status predicted
F;654/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation) #status predicted

Query Match 100.0%; Score 107; DB 1; Length 822;
Best Local Similarity 100.0%; Pred. No. 1.3e-09;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 HSOQAVHKLAKSIPLRRQVTVS 22
|||||
Db 409 HSOQAVHKLAKSIPLRRQVTVS 430
|||||

RESULT 5
S29840
fibroblast growth factor receptor 1 - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 20-Jun-2000
R;Yazaki, N.; Fujita, H.; Ohta, M.; Kawasaki, T.; Itoh, N.
Biochim. Biophys. Acta 1172, 37-42, 1993
A;Title: The structure and expression of the FGF receptor-1 mRNA isoforms in rat tissues
A;Reference number: S29840; MUID:93176824; PMID:8382532
A;Accession: S29840
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-822 <YAZ>
A;Cross-references: EMBL:D12498; NID:g220737; PIDN:BAA02059.1; PID:g220738
C;Superfamily: basic fibroblast growth factor receptor 1; immunoglobulin homology; protei
C;Keywords: ATP; autophosphorylation; growth factor receptor; phosphoprotein; transmembr
F;171-232/Domain: immunoglobulin homology <IMM>
F;476-761/Domain: protein kinase homology <KIN>
F;484-492/Region: protein kinase ATP-binding motif

Query Match 100.0%; Score 107; DB 2; Length 822;
Best Local Similarity 100.0%; Pred. No. 1.3e-09;

Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HSQMAVHKLAKSIPLRRQVTVS 22
 |||||
 Db 409 HSQMAVHKLAKSIPLRRQVTVS 430

RESULT 6
 I49289
 fibroblast growth factor receptor-1, long isoform precursor - mouse
 C:Species: Mus musculus (house mouse)
 C>Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 16-Jul-1999
 C:Accession: I49289
 R:Jin, Y.; Pasumarthi, K.B.; Bock, M.E.; Lytras, A.; Kardami, E.; Cattini, P.A.
 J. Mol. Cell. Cardiol. 26, 1449-1459, 1994
 A:Title: Cloning and expression of fibroblast growth factor receptor-1 isoforms in the mouse
 A:Reference number: I49289; MUID:95205422; PMID:7897669
 A:Accession: I49289
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-822 <RES>
 A:Cross-references: EMBL:U22324; NID:G722340; PIDN:AA52182.1; PID:G722341
 C:Superfamily: basic fibroblast growth factor receptor 1; immunoglobulin homology; protein kinase
 C:Keywords: ATP; growth factor receptor
 F:270-343/Domain: immunoglobulin homology <IMM>
 F:476-761/Domain: protein kinase homology <KIN>
 F:484-492/Region: protein kinase ATP-binding motif

Query Match 100.0%; Score 107; DB 2; Length 822;
 Best Local Similarity 100.0%; Pred. No. 1.3e-09;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HSQMAVHKLAKSIPLRRQVTVS 22
 |||||
 Db 409 HSQMAVHKLAKSIPLRRQVTVS 430

RESULT 7
 JH0393
 fibroblast growth factor receptor precursor - mouse
 N:Contains: protein-tyrosine kinase (EC 2.7.1.112)
 C:Species: Mus musculus (house mouse)
 C>Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 16-Jul-1999
 C:Accession: JH0393
 R:Kouhara, H.; Kasayama, S.; Saito, H.; Matsumoto, K.; Sato, B.
 Biochem. Biophys. Res. Commun. 176, 31-37, 1991
 A:Title: Expression cDNA cloning of fibroblast growth factor (FGF) receptor in mouse brain
 A:Reference number: JH0393; MUID:91207411; PMID:1708247
 A:Accession: JH0393
 A:Molecule type: mRNA
 A:Residues: 1-832 <KOU>
 A:Cross-references: GB:M65053; NID:G193292; PIDN:AAA37620.1; PID:G193293
 A:Experimental source: breast cancer cell line SC-3
 A:Note: the authors translated the codon ATG for residue 397 as Ile and CTG for residue 398
 C:Superfamily: basic fibroblast growth factor receptor 1; immunoglobulin homology; protein kinase
 C:Keywords: ATP; autophosphorylation; duplication; glycoprotein; growth factor receptor;
 F:1-21/Domain: signal sequence #status predicted <SIG>
 F:22-832/Product: fibroblast growth factor receptor #status predicted <FGF>
 F:181-242/Domain: immunoglobulin homology <IMM>
 F:387-409/Domain: transmembrane #status predicted <TM>
 F:486-771/Domain: protein kinase homology <KIN>
 F:494-502/Region: protein kinase ATP-binding motif
 F:67-113,188-240,287-351/Disulfide bonds: #status predicted
 F:524,541,633/Active site: Lys, Glu, Asp #status predicted

Query Match 100.0%; Score 107; DB 2; Length 832;
 Best Local Similarity 100.0%; Pred. No. 1.3e-09;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HSQMAVHKLAKSIPLRRQVTVS 22
 |||||
 Db 419 HSQMAVHKLAKSIPLRRQVTVS 440

RESULT 8
 A39752
 fibroblast growth factor receptor A1 precursor - African clawed frog
 N:Contains: fibroblast growth factor receptor A1, short splice form; protein-tyrosine kinase
 C:Species: Xenopus laevis (African clawed frog)
 C>Date: 07-Feb-1992 #sequence_revision 19-Jan-1996 #text_change 16-Jul-1999
 C:Accession: A39752; B39752
 R:Friesel, R.; Dawid, I.B.
 Mol. Cell. Biol. 11, 2481-2488, 1991
 A:Title: cDNA cloning and developmental expression of fibroblast growth factor receptors
 A:Reference number: A39752; MUID:91203867; PMID:1850097
 A:Accession: A39752
 A:Molecule type: mRNA
 A:Residues: 1-814 <PRI>
 A:Cross-references: GB:M55163; NID:G214893; PIDN:AAA49990.1; PID:G214894
 A:Accession: B39752
 A:Molecule type: mRNA
 A:Residues: 1-30,119-814 <PR2>
 A:Cross-references: GB:M55163
 C:Superfamily: basic fibroblast growth factor receptor 1; autophosphorylation; duplication; glycoprotein; protein kinase
 C:Keywords: alternative splicing; ATP; autophosphorylation; duplication; glycoprotein; protein kinase
 F:1-21/Domain: signal sequence #status predicted <SIG>
 F:22-814/Product: fibroblast growth factor receptor A1, long splice form #status experiment
 F:22-372/Domain: extracellular #status predicted <EXT>
 F:22-30,119-814/Product: fibroblast growth factor receptor A1, short splice form #status predicted <INT>
 F:47-102/Domain: immunoglobulin homology <IM1>
 F:125-132/Region: acidic
 F:167-228/Domain: immunoglobulin homology <IM2>
 F:266-339/Domain: immunoglobulin homology <IM3>
 F:373-393/Domain: transmembrane #status predicted <TM>
 F:394-814/Domain: intracellular #status predicted <INT>
 F:472-757/Domain: protein kinase homology <KIN>
 F:480-488/Region: protein kinase ATP-binding motif
 F:54-100,174-226,273-337/Disulfide bonds: #status predicted
 F:76,116,133,177,223,236,260,292,313,326/Binding site: carbohydrate (Asn) (covalent) #status predicted
 F:510,527,619/Active site: Lys, Glu, Asp #status predicted
 F:624,637/Binding site: magnesium (Asn, Asp) #status predicted
 F:650/Binding site: phosphate -(Tyr) (covalent) (by autophosphorylation) #status predicted

Query Match 90.7%; Score 97; DB 1; Length 814;
 Best Local Similarity 90.9%; Pred. No. 5.8e-08;
 Matches 20; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 HSQMAVHKLAKSIPLRRQVTVS 22
 :||:|||||
 Db 405 NSQMAVHKLAKSIPLRRQVTVS 426

RESULT 9
 TVCHFG
 fibroblast growth factor receptor 1 precursor - chicken
 N:Alternate names: basic fibroblast growth factor receptor
 N:Contains: protein-tyrosine kinase (EC 2.7.1.112) ceki
 C:Species: Gallus gallus (chicken)
 C>Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 16-Jul-1999
 C:Accession: A41345; A33908
 R:Lee, P.L.; Johnson, D.E.; Cousens, L.S.; Fried, V.A.; Williams, L.T.
 Science 245, 57-60, 1989
 A:Title: Purification and complementary DNA cloning of a receptor for basic fibroblast growth factor
 A:Reference number: A41345; MUID:89298406; PMID:2544996
 A:Accession: A41345
 A:Status: nucleic acid sequence not shown; not compared with conceptual translation
 A:Molecule type: mRNA
 A:Residues: 1-819 <LEE>
 A:Note: part of the sequence was confirmed by protein sequencing
 R:Pasquale, E.B.; Singer, S.J.
 Proc. Natl. Acad. Sci. U.S.A. 86, 5449-5453, 1989
 A:Title: Identification of a developmentally regulated protein-tyrosine kinase by using cDNA
 A:Reference number: A33908; MUID:89315814; PMID:2473471
 A:Accession: A33908
 A:Molecule type: mRNA

A;Residues: 1-89,'A','91-685','M',687-819 <PAS>
A;Cross-references: GB:M24637
A;Note: this protein is expressed in embryonic tissues and, at low levels, in adult brain
C;Genetics:
A;Gene: cekl
C;Superfamily: basic fibroblast growth factor receptor 1; immunoglobulin homology; protei
C;Keywords: ATP; autophosphorylation; duplication; glycoprotein; growth factor receptor;
F;1-21/Domain: signal sequence #status predicted <SIG>
F;22-819/Product: fibroblast growth factor receptor 1 #status predicted <MAT>
F;22-374/Domain: extracellular #status predicted <EXT>
F;125-132/Region: acidic
F;159-230/Domain: immunoglobulin homology <IMM>
F;375-395/Domain: transmembrane #status predicted <TMW>
F;396-819/Domain: intracellular #status predicted <INT>
F;474-759/Domain: protein kinase homology <KIN>
F;482-490/Region: protein kinase ATP-binding motif
F;54-100,176-228,275-339/Disulfide bonds: #status predicted
F;76,116,225,238,262,294,315,328/Binding site: carboxydrate (Asn) (covalent) #status pre
F;512,529,621/Active site: Lys, Glu, Asp #status predicted
F;626,639/Binding site: magnesium (Asn, Asp) #status predicted
F;652/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation) #status predi
Query Match 90.7%; Score 97; DB 1; Length 819;
Best Local Similarity 90.9%; Pred. No. 5.9e-08;
Matches 20; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 1 HSQMAVHKLAKSIPLRRQVTVS 22
DB 407 NSQLAVHKLAKSIPLRRQVTVS 428
RESULT 10
A56795
fibroblast growth factor receptor 1 beta-isoform - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 03-Nov-1995 #sequence_revision 03-Nov-1995 #text_change 16-Jul-1999
R;K1m, E.G.; Kwon, H.M.; Burrow, C.R.; Ballermann, B.J.
Am. J. Physiol. 264, F66-F73, 1993
A;Title: Expression of rat fibroblast growth factor receptor 1 as three splicing variant
A;Reference number: A56795; MUID:9318788; PMID:8381605
A;Contents: Sprague-Dawley, kidneys
A;Accession: A56795
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-729 <KIM>
A;Cross-references: GB:S54008; NID:G264804; PIDN:AA854274.1; PID:G264805
A;Note: sequence extracted from NCBI Backbone (NCBIN:124259, NCBIF:124270)
C;Superfamily: basic fibroblast growth factor receptor 1; immunoglobulin homology; protei
C;Keywords: alternative splicing; ATP; growth factor receptor
F;80-141/Domain: immunoglobulin homology <IMM>
F;383-668/Domain: protein kinase homology <KIN>
F;391-399/Region: protein kinase ATP-binding motif
Query Match 88.8%; Score 95; DB 2; Length 729;
Best Local Similarity 95.0%; Pred. No. 1.1e-07;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 HSQMAVHKLAKSIPLRRQVT 20
DB 318 HSQMAVHKLAKSIPLRRQVS 337
RESULT 11
A36477
fibroblast growth factor receptor A2 precursor - African clawed frog
N;Contains: fibroblast growth factor receptor A2, short splice form; protein-tyrosine ki
C;Species: Xenopus laevis (African clawed frog);
C;Date: 08-Mar-1991 #sequence_revision 19-Jan-1996 #text_change 16-Jul-1999
C;Accession: A36477; C39752
R;Musci, T.J.; Amaya, E.; Kirschner, M.W.
Proc. Natl. Acad. Sci. U.S.A. 87, 8365-8369, 1990
A;Title: Regulation of the fibroblast growth factor receptor in early Xenopus embryos.

A;Reference number: A36477; MUID:91045998; PMID:2172985
A;Accession: A36477
A;Molecule type: mRNA
A;Residues: 1-812 <MUS>
A;Cross-references: GB:U24491; GB:M37201; NID:G857677; PIDN:AAA86868.1; PID:G857678
A;Note: 45-Arg, 49-Thr, 50-Val, and 64-Ser also found
R;Priesel, R.; David, I.B. 2481-2488, 1991
Mol. Cell. Biol. 11, 2481-2488, 1991
A;Title: cDNA cloning and developmental expression of fibroblast growth factor receptors
A;Reference number: A39752; MUID:91203867; PMID:1850097
A;Accession: C39752
A;Molecule type: mRNA
A;Residues: 1-30,119-189,'G',191-418,'L',420-636,'R',638-788,'V',790-812 <PRI>
A;Cross-references: GB:M62332; NID:G214899; PIDN:AAA49993.1; PID:G214900
C;Superfamily: basic fibroblast growth factor receptor 1; immunoglobulin homology; protei
C;Keywords: alternative splicing; ATP; autophosphorylation; duplication; glycoprotein; g
protein kinase
F;1-21/Domain: signal sequence #status predicted <SIG>
F;22-812/Product: fibroblast growth factor receptor A2, long splice form #status predi
F;22-372/Domain: extracellular #status predicted <EXT>
F;22-30,119-812/Product: fibroblast growth factor receptor A2, short splice form #status
F;47-102/Domain: immunoglobulin homology <IM1>
F;125-132/Region: acidic
F;167-228/Domain: immunoglobulin homology <IM2>
F;266-339/Domain: immunoglobulin homology <IM3>
F;373-393/Domain: transmembrane #status predicted <TMW>
F;394-812/Domain: intracellular #status predicted <INT>
F;470-755/Domain: protein kinase homology <KIN>
F;478-486/Region: protein kinase ATP-binding motif
F;54-100,174-226,273-337/Disulfide bonds: #status predicted
F;76,116,133,177,223,236,260,292,313,326/Binding site: carboxydrate (Asn) (covalent) #sta
F;508,525,617/Active site: Lys, Glu, Asp #status predicted
F;622,635/Binding site: magnesium (Asn, Asp) #status predicted
F;648/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation) #status predi
Query Match 87.9%; Score 94; DB 1; Length 812;
Best Local Similarity 86.4%; Pred. No. 1.8e-07;
Matches 19; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
QY 1 HSQMAVHKLAKSIPLRRQVTVS 22
DB 405 NSQLAVHKLAKSIPLRRQVTVS 426
RESULT 12
I65223
heparin-binding fibroblast growth factor receptor 2 - rat (fragment)
C;Species: Rattus norvegicus (Norway rat)
C;Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 21-Jul-2000
C;Accession: I65223; I65226
R;Yan, G.; McBride, G.; McKeeshan, W.L.
Biochem. Biophys. Res. Commun. 194, 512-518, 1993
A;Title: Exon skipping causes alteration of the COOH-terminus and deletion of the phosph
A;Reference number: I52281; MUID:93326167; PMID:8333865
A;Accession: I65223
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-415 <RES>
A;Cross-references: GB:L19106; NID:G310144; PIDN:AA802867.1; PID:G310145
A;Accession: I65226
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-415 <RE2>
A;Cross-references: GB:L19109; NID:G310144; PIDN:AA802867.1; PID:G310145
C;Superfamily: basic fibroblast growth factor receptor 1; immunoglobulin homology; protei
C;Keywords: ATP; growth factor receptor; heparin binding
F;73-358/Domain: protein kinase homology <KIN>
F;81-99/Region: protein kinase ATP-binding motif
Query Match 77.6%; Score 83; DB 2; Length 415;
Best Local Similarity 85.7%; Pred. No. 6.4e-06;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 SQMAVHKLAKSIPLRRQVTVS 22
 DB 5 SQPAVHKLTKRIPLRRQVTVS 25

RESULT 13

A54846
 fibroblast growth factor receptor precursor - rat

C:Species: Rattus norvegicus (Norway rat)
 C:Date: 28-Jul-1995 #sequence_revision 28-Jul-1995 #text_change 16-Jul-1999
 C:Accession: A54846
 R:Takagi, Y.; Shrivastava, S.; Miki, T.; Sakaguchi, K.
 J. Biol. Chem. 269, 23743-23749, 1994
 A>Title: Molecular cloning and expression of the acidic fibroblast growth factor receptor
 ability and covalent attachment of heparan sulfate glycosaminoglycan to the receptors.
 A:Reference number: A54846; MUID:94375484; PMID:8089146
 A:Accession: A54846
 A>Status: preliminary; not compared with conceptual translation
 A:Molecule type: mRNA
 A:Residues: 1-707 <TX>
 C:Superfamily: basic fibroblast growth factor receptor 1; immunoglobulin homology; prote
 C:Keywords: ATP; growth factor receptor
 F:57-118/Domain: immunoglobulin homology <IMM>
 F:365-650/Domain: protein kinase homology <KIN>
 F:373-381/Region: protein kinase ATP-binding motif

Query Match 77.6%; Score 83; DB 2; Length 707;
 Best Local Similarity 85.7%; Pred. No. 1.1e-05;
 Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 SQMAVHKLAKSIPLRRQVTVS 22
 DB 297 SQPAVHKLTKRIPLRRQVTVS 317

RESULT 14

A38429
 keratinocyte growth factor receptor precursor - mouse
 C:Species: Mus musculus (house mouse)
 C:Date: 14-Feb-1992 #sequence_revision 14-Feb-1992 #text_change 16-Jul-1999
 C:Accession: A38429
 R:Miki, T.; Fleming, T.P.; Bottaro, D.P.; Rubin, J.S.; Ron, D.; Aaronson, S.A.
 Science 251, 72-75, 1991
 A>Title: Expression cDNA cloning of the KGF receptor by creation of a transforming autoc
 A:Reference number: A38429; MUID:91095977; PMID:1846048
 A:Accession: A38429
 A>Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-707 <MX>
 C:Cross-references: GB:M63503; NID:g198593; PIDN:AAA39377.1; PID:g198594
 C:Superfamily: basic fibroblast growth factor receptor 1; immunoglobulin homology; prote
 C:Keywords: ATP; autophosphorylation; growth factor receptor; phosphoprotein; transmembr
 F:57-118/Domain: immunoglobulin homology <IMM>
 F:365-650/Domain: protein kinase homology <KIN>
 F:373-381/Region: protein kinase ATP-binding motif

Query Match 77.6%; Score 83; DB 2; Length 707;
 Best Local Similarity 85.7%; Pred. No. 1.1e-05;
 Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 SQMAVHKLAKSIPLRRQVTVS 22
 DB 297 SQPAVHKLTKRIPLRRQVTVS 317

RESULT 15

S16236
 fibroblast growth factor receptor precursor - human
 N:Alternate names: bck-related FGF receptor
 C:Species: Homo sapiens (man)
 C:Date: 16-Sep-1992 #sequence_revision 16-Sep-1992 #text_change 16-Jul-1999
 C:Accession: S16236

R:Seno, M.; Sasada, R.; Watanabe, T.; Ishimaru, K.; Igarashi, K.
 Biochim. Biophys. Acta 1089, 244-246, 1991
 A>Title: Two cDNAs encoding novel human FGF receptor.
 A:Reference number: S16236; MUID:91274356; PMID:1647213
 A:Accession: S16236

A:Molecule type: mRNA
 A:Residues: 1-769 <SEN>
 A:Cross-references: EMBL:X56191; NID:g29431; PIDN:CAA39654.1; PID:g29432
 C:Superfamily: basic fibroblast growth factor receptor 1; immunoglobulin homology; prote
 C:Keywords: ATP; autophosphorylation; duplication; glycoprotein; growth factor receptor;
 F:1-21/Domain: signal sequence #status predicted <SIG>
 F:22-769/Product: fibroblast growth factor receptor #status predicted <MAT>
 F:22-378/Domain: extracellular #status predicted <EXT>
 F:132-138/Region: acidic
 F:172-233/Domain: immunoglobulin homology <IMM>
 F:379-399/Domain: transmembrane #status predicted <TM>
 F:400-769/Domain: intracellular #status predicted <INT>
 F:480-765/Domain: protein kinase homology <KIN>
 F:488-496/Region: protein kinase ATP-binding motif
 F:62-107/179-231,278-340/Disulfide bonds: #status predicted
 F:83,123,228,241,265,297,318,329,350/Binding site: carbohydrate (Asn) (covalent) #status
 F:518,535,627/Active site: Lys, Glu, Asp #status predicted

Query Match 77.6%; Score 83; DB 2; Length 769;
 Best Local Similarity 85.7%; Pred. No. 1.2e-05;
 Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 SQMAVHKLAKSIPLRRQVTVS 22
 DB 412 SQPAVHKLTKRIPLRRQVTVS 432

Search completed: August 2, 2004, 09:31:50
 Job time: 4.32895 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 2, 2004, 09:23:08 ; Search time 2.31579 Seconds
(without alignments)
494.666 Million cell updates/sec

Title: US-09-757-415A-3

Perfect score: 107

Sequence: 1 HSCMAVHKLAKSIPLRRQVTS 22

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	107	100.0	822	1	FGRI_HUMAN
2	107	100.0	822	1	FGRI_MOUSE
3	107	100.0	822	1	FGRI_RAT
4	97	90.7	819	1	FGRI_CHICK
5	94	87.9	812	1	FGRI_XENLA
6	83	77.6	821	1	FGRI_HUMAN
7	83	77.6	821	1	FGRI_MOUSE
8	83	77.6	823	1	CEK3_CHICK
9	67	62.6	813	1	FGRI_XENLA
10	47	43.9	4391	1	FGRI_HUMAN
11	45.5	42.5	993	1	SMB2_MOUSE
12	45	42.1	3707	1	FGRI_MOUSE
13	43.5	40.7	135	1	VAL2_TYLCV
14	43	40.2	361	1	AROB_BACCR
15	42	39.3	543	1	EBD2_HUMAN
16	41	38.3	256	1	Y090_MYCTU
17	41	38.3	412	1	CEX6_HUMAN
18	41	38.3	414	1	CEX6_MOUSE
19	41	38.3	717	1	PAL2_ARATH
20	41	38.3	928	1	PM11_CHLPN
21	41	38.3	965	1	AMPN_RABIT
22	40.5	37.9	806	1	FGRI_HUMAN
23	40.5	37.9	989	1	SMB2_MESAU
24	40	37.4	473	1	UXAC_BACST
25	40	37.4	499	1	AMPA_CHLMU
26	40	37.4	705	1	SYN1_HUMAN
27	40	37.4	706	1	SYN1_BOVIN
28	40	37.4	825	1	XLNE_ASPNG
29	40	37.4	1130	1	ABLI_HUMAN
30	39.5	36.9	801	1	FGRI_MOUSE
31	39	36.4	661	1	UAS3_HUMAN
32	39	36.4	1002	1	YEMA_DROME
33	39	36.4	1914	1	STCK_EMENI

ALIGNMENTS

RESULT 1

ID	FGRI_HUMAN	STANDARD;	PRT;	822 AA.
AC	P11362; P17049;			
DT	01-JUL-1989 (Rel. 11, Created)			
DT	01-MAY-1991 (Rel. 18, Last sequence update)			
DT	15-MAR-2004 (Rel. 43, Last annotation update)			
DE	Basic fibroblast growth factor receptor 1 precursor (BC 2.7.1.112)			
DE	(FGFR-1) (bFGF-R) (Fms-like tyrosine kinase-2) (c-fgr).			
GN	FGFR1 OR FLG OR FGFR OR FLT2.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Placenta;			
RX	MEDLINE=90245600; PubMed=2159526;			
RA	Isacchi A., Bergonzoni L., Sarmientos P.,			
RT	"Complete sequence of a human receptor for acidic and basic			
RT	fibroblast growth factors.";			
RL	Nucleic Acids Res. 18:1906-1906(1990).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Neonatal Brain stem;			
RX	MEDLINE=90360977; PubMed=1697263;			
RA	Dionne C.A., Crumley G.R., Bellot F., Kaplow J.M., Searfoss G.,			
RA	Ruta M., Burgess W.H., Jaye M., Schlessinger J.,			
RT	"Cloning and expression of two distinct high-affinity receptors			
RT	cross-reacting with acidic and basic fibroblast growth factors.";			
RL	EMBO J. 9:2685-2692(1990).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RC	MEDLINE=92282615; PubMed=1317750;			
RA	Hattori Y., Odagiri H., Katch O., Sakamoto H., Morita T.,			
RA	Shimotohno K., Tobinai K., Sugimura T., Terada M.,			
RT	"K-sam-related gene, N-sam, encodes fibroblast growth factor receptor			
RT	and is expressed in T-lymphocytic tumors.";			
RL	Cancer Res. 52:3367-3371(1992).			
RN	[4]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Liver;			
RX	MEDLINE=91126480; PubMed=1846977;			
RA	Hou J., Kan M., McKeehan K., McBride G., Adams P., McKeehan W.L.;			
RT	"Fibroblast growth factor receptors from liver vary in three			
RT	structural domains.";			
RL	Science 251:665-668(1991).			
RN	[5]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=92118399; PubMed=1662973;			
RA	Kiefer M.C., Baird A., George-Nascimento C., Nguyen T., Mason O.B.,			
RA	Boley J.J., Valenzuela P., Barr P.J.;			
RT	"Molecular cloning of a human basic fibroblast growth factor receptor			
RT	cDNA and expression of a biologically active extracellular domain in			
RT	a baculovirus system.";			
RL	Growth Factors 5:115-127(1991).			

Q9vm75 drosophila
Q44697 buchnera ap
P55397 rhizobium s
Q8xbj4 escherichia
P37441 salmonella
P33934 caenorhabdi
P33995 escherichia
Q8z3g1 salmonella
Q8z1r8 salmonella
P17163 klebsiella
Q9za87 proteus mir
Q8zb41 yersinia pe

34 39 36.4 2096 1 BP28_DROME
35 38.5 36.0 519 1 TRPE_EUCDN
36 38 35.5 99 1 TRBD_RHISN
37 38 35.5 263 1 UCFA_ECO57
38 38 35.5 263 1 UCFA_SALTY
39 38 35.5 273 1 TCIA_CAEEL
40 38 35.5 284 1 YHBJ_ECOLI
41 38 35.5 284 1 YHBJ_SALTY
42 38 35.5 284 1 YHBJ_SALTY
43 38 35.5 284 1 YPTO_KLEOX
44 38 35.5 284 1 YPTO_PROMI
45 38 35.5 284 1 YZB6_YERPE

16] SEQUENCE FROM N.A.
RP TISSUE=Placenta;
RX MEDLINE=90290512; PubMed=2162671;
RA Itoh N., Terachi T., Ohta M., Seo M.K.;
RT "The complete amino acid sequence of the shorter form of human basic
RT fibroblast growth factor receptor deduced from its cDNA.";
RL Biochem. Biophys. Res. Commun. 169:680-685(1990).
[7]
RP SEQUENCE OF 201-822 FROM N.A.
RA Ruca M., Howk R., Ricca G., Drohan W., Zabelshansky M., Laureys G.,
RA Barton D.B., Francke J., Schlessinger J., Givol D.;
RT "A novel protein tyrosine kinase gene whose expression is modulated
RT during endothelial cell differentiation.";
RL Oncogene 3:9-15(1988).
[8]
RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
RX MEDLINE=90355989; PubMed=2167437;
RA Johnson D.E., Lee P.L., Lu J., Williams L.T.;
RT "Diverse forms of a receptor for acidic and basic fibroblast growth
RT factors.";
RL Mol. Cell. Biol. 10:4728-4736(1990).
[9]
RP ALTERNATIVE SPLICING.
RX MEDLINE=91141499; PubMed=1847500;
RA Guckind S.J., Link D.C., Katamine S., Lacal P., Miki T., Ley T.J.,
RA Robbins K.C.;
RT "A novel c-fgr exon utilized in Epstein-Barr virus-infected B
RT lymphocytes but not in normal monocytes.";
RL Mol. Cell. Biol. 11:1500-1507(1991).
[10]
RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
RC TISSUE=Lung;
RX MEDLINE=91319400; PubMed=1650441;
RA Eisemann A., Ahn J.A., Graziani G., Tronick S.R., Ron D.;
RT "Alternative splicing generates at least five different isoforms of
RT the human basic-FGF receptor.";
RL Oncogene 6:1195-1202(1991).
[11]
RP SEQUENCE FROM N.A.
RP Wennstroem S., Sandstroem C., Claesson-Welsh L.;
RL Submitted (JUL-1990) to the EMBL/GenBank/DBJ databases.
[12]
RP SEQUENCE FROM N.A. (ISOFORM ALPHA B1).
RC TISSUE=Uterus;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Rosinsky S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko V., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalek U., Smallus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
[13]
RP MUTAGENESIS OF TYR-766.
RX MEDLINE=92357144; PubMed=1379697;
RA Peters K.G., Marie J., Wilson E., Ives H.E., Escobedo J.,
RA del Rosario M., Mirda D., Williams L.T.;
RT "Point mutation of an FGF receptor abolishes phosphatidylinositol
RT turnover and Ca²⁺ flux but not mitogenesis.";
RL Nature 358:678-681(1992).
[14]
RP MUTAGENESIS OF TYR-766.
RX MEDLINE=92357145; PubMed=1379698;
RA Mohammadi M., Dionne C.A., Li W., Lin N., Spivak T., Honegger A.M.,
RA Jaye M., Schlessinger J.;
RT "Point mutation in FGF receptor eliminates phosphatidylinositol
RT hydrolysis without affecting mitogenesis.";
RL Nature 358:681-684(1992).
[15]
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 464-762.
RX MEDLINE=96361355; PubMed=8752212;
RA Mohammadi M., McMahon G., Sun L., Tang C., Hirth P., Yeh B.K.,
RA Hubbard S.R., Schlessinger J.;
RT "Structure of the FGF receptor tyrosine kinase domain reveals a novel
RT autoinhibitory mechanism.";
RL Cell 86:577-587(1996).
[16]
RP X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS) OF 464-762.
RX MEDLINE=97284786; PubMed=9139660;
RA Mohammadi M., McMahon G., Sun L., Tang C., Hirth P., Yeh B.K.,
RA Hubbard S.R., Schlessinger J.;
RT "Structures of the tyrosine kinase domain of fibroblast growth factor
RT receptor in complex with inhibitors.";
RL Science 276:955-960(1997).
[17]
RP VARIANT PS ARG-252.
RX MEDLINE=95179173; PubMed=7874169;
RA Muenke M., Schell U., Hehr A., Robin N.H., Losken H.W., Schinzel A.,
RA Puelly L.J., Rutland P., Reardon W., Malcolm S., Winter R.M.;
RT "A common mutation in the fibroblast growth factor receptor 1 gene in
RT Pfeiffer syndrome.";
RL Nat. Genet. 8:269-274(1994).
[18]
RP FUNCTION: Receptor for basic fibroblast growth factor. A shorter
CC form of the receptor could be a receptor for acidic FGF (aFGF).
CC
CC CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
CC tyrosine phosphate.
CC
CC SUBCELLULAR LOCATION: Type I membrane protein.
CC
CC ALTERNATIVE PRODUCTS:
CC Event-Alternative splicing; Named isoforms=7;
CC Comment=Additional isoforms seem to exist;
CC
CC Name=Alpha A1;
CC IsoId=PI1362-1; Sequence=Displayed;
CC Name=Alpha B1;
CC IsoId=PI1362-2; Sequence=VSP_002960;
CC Name=Beta A1;
CC IsoId=PI1362-3; Sequence=VSP_002958;
CC Name=Beta B1;
CC IsoId=PI1362-4; Sequence=VSP_002958, VSP_002960;
CC Name=Gamma A1;
CC IsoId=PI1362-5; Sequence=VSP_002957;
CC Name=Gamma B1;
CC IsoId=PI1362-6; Sequence=VSP_002960;
CC Name=A;
CC IsoId=PI1362-7; Sequence=VSP_002959;
CC
CC DISEASE: Defects in FGFR1 are one of the causes of Pfeiffer
CC syndrome (ACPS) [MIM:101600]; also known as acrocephalosyndactyly
CC type V (ACSV). PS is characterized by craniosynostosis (premature
CC fusion of the skull sutures) with deviation and enlargement of the
CC thumbs and great toes, brachymesophalangy, with phalangal
CC ankylosis and a varying degree of soft tissue syndactyly.
CC
CC DISEASE: Involved in t(8;13)(p12;q12) chromosomal translocation
CC which involves FGFR1 and ZNF198. The resulting transcript is a
CC possible candidate for stem cell leukemia lymphoma syndrome/SCLL.
CC
CC SIMILARITY: Belongs to the fibroblast growth factor receptor
CC family.
CC
CC SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.
CC
CC DATABASE: NAME=Atlas Genet. Cytogenet. Oncol. Haematol.;
CC WWW="http://www.infobiogen.fr/services/chronocancer/genes/FGFR1113.html".

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

Query Match 100.0%; Score 107; DB 1; Length 822;
 Best Local Similarity 100.0%; Pred. NO. 2.1e-10;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 H5QMAVHKAKSIPLRQVTVS 22
 |||||
 Db 409 H5QMAVHKAKSIPLRQVTVS 430

RESULT 2

FGFR1_MOUSE STANDARD; PRT; 822 AA.
 AC P16092; Q01736; Q61562;
 DT 01-APR-1990 (Rel. 14, Created)
 DT 01-MAY-1991 (Rel. 18, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Basic fibroblast growth factor receptor 1 precursor (EC 2.7.1.112)
 DE (FGFR-1) (bFGF-R) (MFR).
 GN FGFR1 OR FLG.
 OS Mus musculus (Mouse).
 CC Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Rutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=90160373; PubMed=1699490;
 RA "Reid H.H., Wilks A.F., Bernard O.;
 RT "Two forms of the basic fibroblast growth factor receptor-like mRNA
 RT are expressed in the developing mouse brain.";
 RL Proc. Natl. Acad. Sci. U.S.A. 87:1596-1600(1990).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BALE/c; TISSUE=Brain;
 RX MEDLINE=90285603; PubMed=2161096;
 RA Safran A., Avioli A., Orr-Urtreger A., Neufeld G., Lonai P.,
 RA Givol D., Yarden Y.;
 RT "The murine flg gene encodes a receptor for fibroblast growth
 RT factor.";
 RL Oncogene 5:635-643(1990).
 RN [3]
 RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
 RX MEDLINE=91207411; PubMed=1708247;
 RA Kouhara H., Kasayama S., Saito H., Matsumoto K., Sato B.;
 RT "Expression cDNA cloning of fibroblast growth factor (FGF) receptor
 RT in mouse breast cancer cells: a variant form in FGF-responsive
 RT transformed cells.";
 RL Biochem. Biophys. Res. Commun. 176:31-37(1991).
 RN [4]
 RP SEQUENCE FROM N.A. (ISOFORM 2).
 RX MEDLINE=90272715; PubMed=2161540;
 RA Mansukhani A., Moscatelli D., Talarico D., Levitska V., Basilico C.;
 RT "A murine fibroblast growth factor (FGF) receptor expressed in CHO
 RT cells is activated by basic FGF and Kaposi FGF";
 RL Proc. Natl. Acad. Sci. U.S.A. 87:4378-4382(1990).
 RN [5]
 RP SEQUENCE OF 1-15 FROM N.A., AND ALTERNATIVE SPLICING.
 RX MEDLINE=95100926; PubMed=7802632;
 RA Harada T., Saito H., Kouhara H., Kurebayashi S., Kasayama S.,
 RA Terakawa N., Kishimoto T., Sato B.;
 RT "Murine fibroblast growth factor receptor 1 gene generates multiple
 RT messenger RNAs containing two open reading frames via alternative
 RT splicing.";
 RL Biochem. Biophys. Res. Commun. 205:1057-1063(1994).
 CC -1- FUNCTION: Receptor for basic fibroblast growth factor. A shorter
 CC form of the receptor could be a receptor for acidic FGF (aFGF).
 CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
 CC tyrosine phosphate.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event-Alternative splicing; Named isoforms=3;
 CC Name=1;

CC IsoId=P16092-1; Sequence=Displayed;
 CC Name=2;
 CC IsoId=P16092-2; Sequence=VSP_002962;
 CC Name=3; Synonyms=Variant;
 CC IsoId=P16092-3; Sequence=VSP_002961, VSP_002963;
 CC -1- SIMILARITY: Belongs to the fibroblast growth factor receptor
 CC family.
 CC -1- SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; M28998; AAA37290.1; -;
 CC EMBL; X51893; CAA36175.1; -;
 CC EMBL; M65053; AAA37620.1; -;
 CC EMBL; M33760; AAA37622.1; -;
 CC EMBL; S74765; AAB32845.1; ALT_SEQ.
 CC FIR; A34849; TVMSFG.
 CC HSSP; P11362; IFGK.
 CC MGD; MGI:95522; Fgf1.
 CC GO; GO:0007420; P:brain development; IMP.
 CC GO; GO:0042472; P:inner ear morphogenesis; IMP.
 CC GO; GO:0007435; P:salivary gland morphogenesis; IMP.
 CC InterPro; IPR007110; IG-like.
 CC InterPro; IPR003598; IG_C2.
 CC InterPro; IPR000719; Prot_kinase.
 CC InterPro; IPR001245; Tyr_kinase.
 CC InterPro; IPR008266; Tyr_kinase_AS.
 CC Pfam; PF00047; IG_3.
 CC Pfam; PF00069; pkinase_1.
 CC PRINTS; PR00109; TYRKINASE.
 CC ProDom; PD000001; Prot_kinase; 1.
 CC SMART; SM00408; IGC2; 3.
 CC SMART; SM00219; TyrKc; 1.
 CC PROSITE; PS00835; IG LIKE; 3.
 CC PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 CC PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
 CC PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
 CC Receptor; Glycoprotein; Tyrosine-protein kinase; ATP-binding;
 CC Transferrase; Phosphorylation; Transmembrane; Signal;
 CC Immunoglobulin domain; Repeat; Alternative splicing.
 CC SIGNAL 1 21
 CC CHAIN 22 822
 CC 1. BASIC FIBROBLAST GROWTH FACTOR RECEPTOR
 CC 1. EXTRACELLULAR (POTENTIAL).
 CC POTENTIAL.
 CC CYTOPLASMIC (POTENTIAL).
 CC IG-LIKE C2-TYPE 1.
 CC IG-LIKE C2-TYPE 2.
 CC IG-LIKE C2-TYPE 3.
 CC PROTEIN KINASE
 CC ATP (BY SIMILARITY).
 CC ATP (BY SIMILARITY).
 CC BY SIMILARITY.
 CC PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
 CC POTENTIAL.
 CC POTENTIAL.
 CC N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC Q -> QGSSSWPLWAAA (in isoform 3).
 CC /FTid=VSP_002961.


```

RP SEQUENCE FROM N.A.
RX MEDLINE=89298406; PubMed=2544996;
RA Lee P.L., Johnson D.E., Cousens L.S., Fried V.A., Williams L.T.;
RT "Purification and complementary DNA cloning of a receptor for basic
RT fibroblast growth factor.";
RL Science 245:57-60(1989).
CC -!- FUNCTION: RECEPTOR FOR BASIC FIBROBLAST GROWTH FACTOR.
CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
CC tyrosine phosphate.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- SIMILARITY: Belongs to the fibroblast growth factor receptor
CC family.
CC -!- SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M24637; AAA48663.1; -.
CC PIR; A41345; TVCHFG.
CC HSSP; P11362; 1FGK.
CC InterPro; IPR007110; Ig-like.
CC InterPro; IPR003598; Ig_c2.
CC InterPro; IPR00719; Prot_kinase.
CC InterPro; IPR001245; Tyr_pkinase.
CC InterPro; IPR008266; Tyr_pkinase_AS.
CC Pfam; PF00047; Ig; 3.
CC Pfam; PF00069; pkinase; 1.
CC PRINTS; PR00109; TYRKINASE.
CC ProDom; PD000001; Prot_kinase; 1.
CC SMART; SM00408; IGC2; 3.
CC SMART; SM00219; TyRK; 1.
CC PROSITE; PS50835; IG-LIKE; 3.
CC PROSITE; PS00107; PROTEIN KINASE ATP; 1.
CC PROSITE; PS50011; PROTEIN KINASE DOM; 1.
CC PROSITE; PS00109; PROTEIN KINASE TYR; 1.
CC Receptor; Glycoprotein; Tyrosine-protein kinase; ATP-binding;
CC Transferase; Phosphorylation; Transmembrane; Immunoglobulin domain;
CC Repeat; Signal.
CC SIGNAL 1 21
CC CHAIN 22 819
CC
CC BASIC FIBROBLAST GROWTH FACTOR RECEPTOR
CC
CC 1.
CC EXTRACELLULAR (POTENTIAL).
CC POTENTIAL.
CC CYTOPLASMIC (POTENTIAL).
CC IG-LIKE C2-TYPE 1.
CC IG-LIKE C2-TYPE 2.
CC IG-LIKE C2-TYPE 3.
CC ASP/GLU-RICH (HIGHLY ACIDIC).
CC PROTEIN KINASE.
CC NP_BIND 482 490
CC ATP (BY SIMILARITY).
CC ACT_SITE 512 512
CC BY SIMILARITY.
CC MOD_RES 652 652
CC PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
CC POTENTIAL.
CC POTENTIAL.
CC N-LINKED (GLNAC. .) (POTENTIAL).
CC N-LINKED (GLNAC. .) (POTENTIAL).
CC N-LINKED (GLNAC. .) (POTENTIAL).
CC N-LINKED (GLNAC. .) (POTENTIAL).
CC N-LINKED (GLNAC. .) (POTENTIAL).
CC N-LINKED (GLNAC. .) (POTENTIAL).
CC N-LINKED (GLNAC. .) (POTENTIAL).
CC N-LINKED (GLNAC. .) (POTENTIAL).
CC A -> R (IN REF. 3).
CC CONFLICT 819 AA; 91576 MW; 7E030B7AE5181DDC CRC64;
SQ SEQUENCE 819 AA; 91576 MW; 7E030B7AE5181DDC CRC64;

```

Query Match

90.7%; Score 97; DB 1; Length 819;

```

Best Local Similarity 90.9%; Pred. No. 1.le-08;
Matches 20; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
OY 1 HSMQAVHKLAKSIPLRQQTVS 22
Db 407 NSQLAVHKLAKSIPLRQQTVS 428
RESULT 5
FGRI_XENLA
ID FGRI_XENLA STANDARD; PRT; 812 AA.
AC P22182;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Fibroblast growth factor receptor 1 precursor (EC 2.7.1.112).
GN FGFR-1.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN RP
RP SEQUENCE FROM N.A.
RX MEDLINE=91045996; PubMed=2172985;
RA Musci T.J., Amaya E., Kirschner M.W.;
RT "Regulation of the fibroblast growth factor receptor in early Xenopus
RT embryos.";
RL Proc. Natl. Acad. Sci. U.S.A. 87:8365-8369(1990).
CC -!- FUNCTION: RECEPTOR FOR BASIC FIBROBLAST GROWTH FACTOR.
CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
CC tyrosine phosphate.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- SIMILARITY: Belongs to the fibroblast growth factor receptor
CC family.
CC -!- SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U24491; AAA86868.1; -.
CC PIR; A36477; A36477.
CC HSSP; P11362; 1FGK.
CC InterPro; IPR007110; Ig-like.
CC InterPro; IPR003598; Ig_c2.
CC InterPro; IPR000719; Prot_kinase.
CC InterPro; IPR001245; Tyr_pkinase.
CC InterPro; IPR008266; Tyr_pkinase_AS.
CC Pfam; PF00047; Ig; 3.
CC Pfam; PF00069; pkinase; 1.
CC PRINTS; PR00109; TYRKINASE.
CC ProDom; PD000001; Prot_kinase; 1.
CC SMART; SM00408; IGC2; 3.
CC SMART; SM00219; TyRK; 1.
CC PROSITE; PS50835; IG-LIKE; 3.
CC PROSITE; PS00107; PROTEIN KINASE ATP; 1.
CC PROSITE; PS50011; PROTEIN KINASE DOM; 1.
CC PROSITE; PS00109; PROTEIN KINASE TYR; 1.
CC Receptor; Glycoprotein; Tyrosine-protein kinase; ATP-binding;
CC Transferase; Phosphorylation; Transmembrane; Immunoglobulin domain;
CC Repeat; Signal.
CC SIGNAL 1 20
CC CHAIN 21 812
CC DOMAIN 21 371
CC TRANSMEM 372 393
CC DOMAIN 394 812
CC DOMAIN 33 110
CC DOMAIN 154 242
CC
CC BY SIMILARITY.
CC FIBROBLAST GROWTH FACTOR RECEPTOR 1.
CC EXTRACELLULAR (POTENTIAL).
CC POTENTIAL.
CC CYTOPLASMIC (POTENTIAL).
CC IG-LIKE C2-TYPE 1.
CC IG-LIKE C2-TYPE 2.

```

FT DOMAIN 251 353 IG-LIKE C2-TYPE 3.
 FT DOMAIN 472 761 PROTEIN KINASE.
 FT NP_BIND 478 486 ATP (BY SIMILARITY).
 FT BINDING 508 508 ATP (BY SIMILARITY).
 FT ACT_SITE 617 617 BY SIMILARITY.
 FT MOD_RES 648 648 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
 FT CARBOHYD 76 76 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 116 116 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 133 133 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 177 177 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 223 223 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 236 236 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 260 260 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 292 292 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 313 313 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 326 326 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 812 AA; 90502 MW; B06333BAFEAC5C9B CRC64;
 Query March 87.9%; Score 94; DB 1; Length 812;
 Best Local Similarity 86.4%; Pred. No. 3.6e-08;
 Matches 19; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
 Qv 1 H5QMAVHKLAKSLPRRQVTVS 22
 Db 405 NSQLAVHKLAKSLPRRQVTVS 426
 RESULT-6
 FGR2 HUMAN
 ID FGR2 HUMAN STANDARD; PRT; 821 AA.
 AC P21802; P18443; Q01742; Q12922; Q14300; Q14301; Q14302; Q14303;
 AC Q14304; Q14305; Q14672; Q14718; Q96K19; Q96K30; Q96K31;
 AC Q96K32; Q9NZU2; Q9NZU3; Q9UD01; Q9UD02; Q9UIH3; Q9UIH4; Q9UIH5;
 AC Q9UIH6; Q9UIH7; Q9UIH8; Q9UIH9; Q9UMC6; Q9UMC7; Q9UOH7; Q9UOH8;
 AC Q9UOH9; Q9UQ10;
 DT 01-NOV-1990 (Rel. 16, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Fibroblast growth factor receptor 2 precursor (EC 2.7.1.112) (FGFR-2)
 DE (keratinocyte growth factor receptor 2).
 GN FGFR2 OR BEK OR KSAM.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RC TISSUE=Neonatal brain stem;
 RX MEDLINE=90360977; PubMed=1697263;
 RA Dionne C.A., Crumley G.R., Bellot F., Kaplow J.M., Searfoss G.,
 RA Ruta M., Burgess W.H., Jaye M., Schlesinger J.;
 RT "Cloning and expression of two distinct high-affinity receptors
 RT cross-reacting with acidic and basic fibroblast growth factors";
 RL ENBO J. 9:2685-2692(1990).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORM 16).
 RX MEDLINE=91045961; PubMed=2172978;
 RA Housaint E., Blanquet P.R., Champion-Arnaud P., Gesnel M.C.,
 RA Torriglia A., Courtois Y., Breathnach R.;
 RT "Related fibroblast growth factor receptor genes exist in the human
 RT genome";
 RL Proc. Natl. Acad. Sci. U.S.A. 87:8180-8184(1990).
 RN [3]
 RP SEQUENCE FROM N.A. (ISOFORM 17).
 RX MEDLINE=91274356; PubMed=1647213;
 RA Seno M., Sasada R., Watanabe T., Ishimaru K.;
 RT "Two cDNAs encoding novel human FGF receptor";
 RL Biochim. Biophys. Acta 1089:244-246(1991).
 RN [4]
 RP SEQUENCE FROM N.A. (ISOFORM 4).
 RC TISSUE=Stomach cancer;
 RX MEDLINE=90332706; PubMed=2377625;
 RA Hattori Y., Odagiri H., Nakatani H., Miyagawa K., Naito K.,

RA Sakamoto H., Katoh O., Yoshida T., Sugimura T., Terada M.;
 RT "K-sam, an amplified gene in stomach cancer, is a member of the
 RT heparin-binding growth factor receptor genes";
 RL Proc. Natl. Acad. Sci. U.S.A. 87:5983-5987(1990).
 RN [5]
 RP SEQUENCE FROM N.A. (ISOFORMS 5; 14 AND 15).
 RX MEDLINE=92212948; PubMed=1313574;
 RA Katoh M., Hattori Y., Sasaki H., Tanaka M., Sugano K., Yazaki Y.,
 RA Sugimura T., Terada M.;
 RT "K-sam gene encodes secreted as well as transmembrane receptor
 RT tyrosine kinase";
 RL Proc. Natl. Acad. Sci. U.S.A. 89:2960-2964(1992).
 RN [6]
 RP SEQUENCE FROM N.A. (ISOFORM 3).
 RC TISSUE=Placenta;
 RX MEDLINE=93016048; PubMed=1400433;
 RA Dell K.R., Williams L.T.;
 RT "A novel form of fibroblast growth factor receptor 2. Alternative
 RT splicing of the third immunoglobulin-like domain confers ligand
 RT binding specificity";
 RL J. Biol. Chem. 267:21225-21229(1992).
 RN [7]
 RP SEQUENCE FROM N.A. (ISOFORMS 3 AND 19), AND VARIANT ARG-613.
 RC TISSUE=Mammary Gland;
 RX MEDLINE=92108030; PubMed=1309608;
 RA Niki T., Sottaro D.P., Fleming T.P., Smith C.L., Burgess W.H.,
 RA Chan A.M.-L., Aaronson S.A.;
 RT "Determination of ligand-binding specificity by alternative splicing:
 RT two distinct growth factor receptors encoded by a single gene";
 RL Proc. Natl. Acad. Sci. U.S.A. 89:246-250(1992).
 RN [8]
 RP SEQUENCE FROM N.A. (ISOFORM 19).
 RC TISSUE=Breast, and Cornea;
 RX MEDLINE=95170769; PubMed=7866434;
 RA Wilson S.E., Weng J., Chwang E.L., Gollahan L., Leitch A.M.,
 RA Shay J.W.;
 RT "Hepatocyte growth factor (HGF), keratinocyte growth factor (KGF), and
 RT their receptors in human breast cells and tissues: alternative
 RT receptors";
 RL Cell. Mol. Biol. Res. 40:337-350(1994).
 RN [9]
 RP ERRATUM.
 RC TISSUE=Breast, and Cornea;
 RA Wilson S.E., Weng J., Chwang E.L., Gollahan L., Leitch A.M.,
 RA Shay J.W.;
 RL Cell. Mol. Biol. Res. 40:707-707(1994).
 RN [10]
 RP SEQUENCE FROM N.A. (ISOFORM 1), AND VARIANT CS SER-342.
 RC TISSUE=Blood;
 RA Steinberger D., Mueller U.;
 RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
 RN [11]
 RP SEQUENCE FROM N.A. (ISOFORMS 7; 9; 10; 11; 12 AND 13), AND VARIANT
 RP ARG-613.
 RX MEDLINE=20090220; PubMed=10626794;
 RA Ueda T., Sasaki H., Kuwahara Y., Nezu M., Shibuya T., Sakamoto H.,
 RA Ishii H., Yanagihara K., Mafune K.-I., Makuuchi M., Terada M.;
 RT "Deletion of the carboxyl-terminal exons of K-sam/FGFR2 by short
 RT homology-mediated recombination, generating preferential expression
 RT of specific messenger RNAs";
 RL Cancer Res. 59:6080-6086(1999).
 RN [12]
 RP SEQUENCE FROM N.A. (ISOFORMS 5; 6; 8; 14 AND 18).
 RX MEDLINE=21845873; PubMed=11856867;
 RA Ingersoll R.G., Paznekas W.A., Tran A.K., Scott A.F., Jiang G.,
 RA Jabs E.W.;
 RT "Fibroblast growth factor receptor 2 (FGFR2): genomic sequence and
 RT variations";
 RL Cytogenet. Cell Genet. 94:121-126(2001).
 RN [13]
 RP SEQUENCE FROM N.A. (ISOFORM 3).
 RA Lind D.L., Cox D.R.;
 RT "Sequence and polymorphisms in fibroblast growth factor receptor 2

RT (FGFR2) gene in humans.";
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
RN [14]
RP SEQUENCE OF 314-427 FROM N.A.
RX MEDLINE=20177482; PubMed=10712195;
RA Glaser R.L., Jiang W., Boyadjiev S.A., Tran A.K., Zachary A.A.,
RA Van Maldergem L., Johnson D., Walsh S., Oldridge M., Wall S.A.,
RA Wilkie A.O.M., Jabs E.W.;
RT "Paternal origin of FGFR2 mutations in sporadic cases of Crouzon
RT syndrome and Pfeiffer syndrome.";
RL Am. J. Hum. Genet. 56:768-777(2000).
RN [15]
RP SEQUENCE OF 1-209; 212-767 AND 771-821 FROM N.A. (ISOFORMS 5; 14 AND
RP 18).
RX MEDLINE=99214070; PubMed=10196476;
RA Zhang Y., Gorry M.C., Post J.C., Ehrlich G.D.;
RT "Genomic organization of the human fibroblast growth factor receptor 2
RT (FGFR2) gene and comparative analysis of the human FGFR gene family.";
RL Gene 230:69-79(1999).
RN [16]
RP SEQUENCE OF 249-313 FROM N.A., AND VARIANTS AS TRP-252 AND ARG-253.
RX MEDLINE=95397814; PubMed=7668257;
RA Park W.-J., Theda C., Maestri N.E., Meyers G.A., Fryburg J.S.,
RA Dufresne C., Cohen M.M. Jr., Jabs E.W.;
RT "Analysis of phenotypic features and FGFR2 mutations in Apert
RT syndrome.";
RL Am. J. Hum. Genet. 57:321-328(1995).
RN [17]
RP SEQUENCE OF 251-259 FROM N.A.
RX MEDLINE=96253074; PubMed=8676562;
RA Wada C., Ishigaki M., Toyooka Y., Yamabe H., Ohnuki Y., Takada F.,
RA Yamazaki Y., Ohtani H.;
RT "Nucleotide sequences at intron 6 and exon 7 junction of fibroblast
RT growth factor receptor 2 and rapid mutational analysis in Apert
RT syndrome.";
RL Rinsho Byori 44:435-438(1996).
RN [18]
RP SEQUENCE OF 251-318 FROM N.A.
RX MEDLINE=96241572; PubMed=8673103;
RA Moloney D.M., Slaney S.F., Oldridge M., Wall S.A., Sahlin P.,
RA Stenman G., Wilkie A.O.M.;
RT "Exclusive paternal origin of new mutations in Apert syndrome.";
RL Nat. Genet. 13:48-53(1996).
RN [19]
RP SEQUENCE OF 263-361 FROM N.A., AND VARIANTS CS PRO-289; ARG-338;
RP SER-342; TYR-342; GLY-344 AND CYS-354.
RX MEDLINE=96090259; PubMed=7581378;
RA Gorry M.C., Preston R.A., White G.J., Zhang Y., Singhal V.K.,
RA Losken H.W., Parker M.G., Nwokoro N.A., Post J.C., Ehrlich G.D.;
RT "Crouzon syndrome: mutations in two spliceforms of FGFR2 and a common
RT point mutation shared with Jackson-Weiss syndrome.";
RL Hum. Mol. Genet. 4:1387-1390(1995).
RN [20]
RP VARIANTS CS HIS-340; ARG-342; SER-342; TYR-342 AND CYS-354.
RX MEDLINE=95078932; PubMed=7987400;
RA Reardon W., Winter R.M., Rutland P., Pulleyn L.J., Jones B.M.,
RA Malcolm S.;
RT "Mutations in the fibroblast growth factor receptor 2 gene cause
RT Crouzon syndrome.";
RL Nat. Genet. 8:98-103(1994).
RN [21]
RP VARIANTS CS CYS-328 AND CYS-347, AND VARIANT JWS GLY-344.
RX MEDLINE=95179174; PubMed=7874170;
RA Jabs E.W., Li X., Scott A.F., Meyers G.A., Chen W., Eccles M., Mao J.,
RA Charnas L.R., Jackson C.E., Jaye M.;
RT "Jackson-Weiss and Crouzon syndromes are allelic with mutations in
RT fibroblast growth factor receptor 2.";
RL Nat. Genet. 8:275-279(1994).
RN [22]
RP VARIANTS CS
RX MEDLINE=95384152; PubMed=7655462;
RA Oldridge M., Wilkie A.O.M., Slaney S.F., Poole M.D., Pulleyn L.J.,
RA Rutland P., Hockley A.D., Wake M.J.C., Goldin J.H., Winter R.M.,

RA Reardon W., Malcolm S.;
RT "Mutations in the third immunoglobulin domain of the fibroblast growth
RT factor receptor-2 gene in Crouzon syndrome.";
RL Hum. Mol. Genet. 4:1077-1082(1995).
RN [23]
RP VARIANTS CS GLY-290; TRP-342 AND CYS-354, AND VARIANT JWS ARG-342.
RX MEDLINE=96133301; PubMed=8528214;
RA Park W.-J., Meyers G.A., Li X., Theda C., Day D., Orlow S.J.,
RA Jones M.C., Jabs E.W.;
RT "Novel FGFR2 mutations in Crouzon and Jackson-Weiss syndromes show
RT Query Match 77.6%; Score 83; DB 1; Length 821;
Best Local Similarity 85.7%; Pred. No. 2.9e-06;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 2 SQMAVHKLAKSIPLRRQTVTS 22
DB 411 SQPAYHKLTKRIPLRRQTVTS 431
RESULT 7
FGFR2_MOUSE
ID FGFR2_MOUSE STANDARD; PRT; 821 AA.
AC P21803; O55141; Q00389; Q61342;
DT 01-NOV-1991 (Rel. 18, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Fibroblast growth factor receptor 2 precursor (EC 2.7.1.112) (FGFR-2)
DE (keratinocyte growth factor receptor).
GN FGFR2 OR ECT1 OR BEK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM LONG).
RC TISSUE=Brain, and Liver;
RX MEDLINE=92228773; PubMed=1373495;
RA Mansukhani A., Dell'Era P., Moscatelli D., Kornbluth S.,
RA Hanafusa H., Basilico C.;
RT "Characterization of the murine BEK fibroblast growth factor (FGF)
RT receptor: activation by three members of the FGF family and
RT requirement for heparin.";
RL Proc. Natl. Acad. Sci. U.S.A. 89:3305-3309(1992).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM LONG).
RC STRAIN=BALE/c; TISSUE=Brain;
RX MEDLINE=91270892; PubMed=1711190;
RA Raz V., Kelman Z., Avioli A., Neufeld G., Givol D., Yarden Y.;
RT "PCR-based identification of new receptors: molecular cloning of a
RT receptor for fibroblast growth factors.";
RL Oncogene 6:753-760(1991).
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM SHORT).
RX MEDLINE=91095977; PubMed=1846048;
RA Miki T., Fleming T.P., Bottaro D.P., Rubin J.S., Ron D.,
RA Aaronson S.A.;
RT "Expression cDNA cloning of the KGF receptor by creation of a
RT transforming autoclone loop.";
RL Science 251:72-75(1991).
RN [4]
RP SEQUENCE FROM N.A. (ISOFORM LONG).
RX MEDLINE=98167854; PubMed=9499422;
RA Twigg S.R.F., Burns H.D., Oldridge M., Heath J.K., Wilkie A.O.M.;
RT "Conserved use of non-canonical 5' splice site (/GA) in alternative
RT splicing by fibroblast growth factor receptors 1, 2 and 3.";
RL Hum. Mol. Genet. 7:685-691(1998).
RN [5]
RP SEQUENCE OF 477-821 FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=89219016; PubMed=2468999;
RA Kornbluth S., Paulson K.E., Hanafusa H.;
RT "Novel tyrosine kinase identified by phosphotyrosine antibody

screening of cDNA libraries.";
 Mol. Cell. Biol. 8:5541-5544(1988).
 CC -!- FUNCTION: RECEPTOR FOR ACIDIC AND BASIC FIBROBLAST GROWTH FACTORS.
 CC POSSESSES A HIGHER AFFINITY FOR ACIDIC THAN FOR BASIC PGF'S.
 CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
 CC tyrosine phosphate.
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -!- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Name=Long;
 CC IsoId=P211803-1; Sequence=Displayed;
 CC Name=Short;
 CC IsoId=P211803-2; Sequence=VSP_002985, VSP_002986, VSP_002987;
 CC -!- SIMILARITY: Belongs to the fibroblast growth factor receptor
 CC family.
 CC -!- SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M86441; AA37286.1; -;
 DR EMBL; X55441; CAA39083.1; -;
 DR EMBL; M63503; AAA39377.1; -;
 DR EMBL; Y16152; CAA76098.1; -;
 DR EMBL; Y16167; CAA76099.1; -;
 DR EMBL; M23362; AAA37285.1; -;
 DR PIR; A44424; TVMSBK.
 DR PIR; S17295; S17295.
 DR HSSP; P11362; IFKC.
 DR MGD; MGI:95523; Fgfr2.
 DR GO; GO:0007435; P:salivary gland morphogenesis; IMP.
 DR InterPro; IPR007110; Ig-like.
 DR InterPro; IPR003598; Ig_C2.
 DR InterPro; IPR000719; Prot_kinase.
 DR InterPro; IPR001245; Tyr_kinase.
 DR InterPro; IPR008266; Tyr_kinase_AS.
 DR Pfam; PF00047; Ig; 3.
 DR Pfam; PF00069; pkinase; 1.
 DR PRINTS; PR00109; TYRKINASE.
 DR ProDom; PD000001; Prot_kinase; 1.
 DR SMART; SM00408; IGC2; 3.
 DR SMART; SM00219; TyrKc; 1.
 DR PROSITE; PS50835; IG_LIKE; 3.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
 KW Receptor; Glycoprotein; Tyrosine-protein kinase; ATP-binding;
 KW Transferase; Phosphorylation; Transmembrane; Signal;
 KW Immunoglobulin domain; Repeat; Alternative splicing.
 FT SIGNAL 1 21
 FT CHAIN 22 821 FIBROBLAST GROWTH FACTOR RECEPTOR 2.
 FT DOMAIN 22 377 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 378 398 POTENTIAL.
 FT DOMAIN 399 821 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 39 125 IG-LIKE C2-TYPE 1.
 FT DOMAIN 154 247 IG-LIKE C2-TYPE 2.
 FT DOMAIN 256 358 IG-LIKE C2-TYPE 3.
 FT DOMAIN 481 770 PROTEIN KINASE.
 FT NP_BIND 487 495 ATP (BY SIMILARITY).
 FT BINDING 517 517 ATP (BY SIMILARITY).
 FT ACT_SITE 626 626 BY SIMILARITY.
 FT MOD_RES 657 657 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
 FT DISULFID 62 107 POTENTIAL.
 FT DISULFID 179 231 POTENTIAL.
 FT DISULFID 278 342 POTENTIAL.
 FT CARBOHYD 83 83 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 123 123 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 147 147 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 228 228 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 241 241 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 265 265 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 297 297 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 318 318 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 331 331 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT VARSPLIC 37 37 E -> G (in isoform Short).
 FT VARSPLIC 38 152 Missing (in isoform Short).
 FT VARSPLIC 38 152 /FTid=VSP_002985.
 FT VARSPLIC 314 361 /FTid=VSP_002986.
 FT AAGVNTTDETEVLVIRNVTFEDAGEYTCIAGNSIGISFHS
 FT AMLVLP -> HSGINSNAEVLALFNVTENDAGEYICKVS
 FT NYIGANOSAWLTVLPQQ (in isoform Short).
 FT /FTid=VSP_002987.
 FT A -> V (IN REF. 2).
 FT GE -> RG (IN REF. 2).
 FT E -> R (IN REF. 2).
 FT I -> Y (IN REF. 2).
 FT DV -> R (IN REF. 2).
 FT C -> V (IN REF. 2 AND 3).
 FT S -> P (IN REF. 2 AND 3).
 FT W -> R (IN REF. 2).
 FT Y -> I (IN REF. 2).
 FT E -> R (IN REF. 2).
 FT N -> Y (IN REF. 2).
 FT SEQUENCE 821 AA; 91983 MW; FCD828ADD61F4414 CRC64;
 Query Match 77.6%; Score 83; DB 1; Length 821;
 Best Local Similarity 85.7%; Pred. NO. 2.9e-06;
 Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 2 SQMAVHKLAKSIPLERQVTVS 22
 DB 411 SQPAVHKLTKRPLRQVTVS 431
 RESULT 8
 CEK3 CHICK
 ID_CEK3 CHICK STANDARD; PRT; 823 AA.
 AC P18461;
 DT 01-NOV-1990 (Rel. 16, Created)
 DT 01-NOV-1990 (Rel. 16, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Tyrosine kinase receptor CEK3 precursor (EC 2.7.1.112).
 GN CEK3.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=9032672; PubMed=2165604;
 RA Pasquale E.B.;
 RT "A distinctive family of embryonic protein-tyrosine kinase
 RT receptors".
 RL Proc. Natl. Acad. Sci. U.S.A. 87:5812-5816(1990).
 CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
 CC tyrosine phosphate.
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -!- SIMILARITY: Belongs to the fibroblast growth factor receptor
 CC family.
 CC -!- SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M35196; AAA48665.1; -;


```

DR PIR; B35963; B35963.
DR HSP; P11362; LFQK.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003598; Ig_c2.
DR InterPro; IPR00719; Prot_kinase.
DR InterPro; IPR001245; Tyr_pkinase.
DR InterPro; IPR008266; Tyr_pkinase_AS.
DR Pfam; PF00047; Ig; 3.
DR Pfam; PF00069; pkinase; 1.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00408; IGC2; 3.
DR SMART; SM00219; TYRK; 1.
DR PROSITE; PS00835; IG LIKE; 3.
DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
DR PROSITE; PS00011; PROTEIN KINASE DOM; 1.
DR PROSITE; PS00109; PROTEIN KINASE TYR; 1.
KW Receptor; Glycoprotein; Tyrosine-protein kinase; ATP-binding;
KW Transferase; Phosphorylation; Transmembrane; Immunoglobulin domain;
KW Repeat; Signal.
FT SIGNAL 1 23
FT CHAIN 24 823
FT DOMAIN 24 379
FT TRANSMEM 380 400
FT DOMAIN 401 823
FT DOMAIN 42 128
FT DOMAIN 156 249
FT DOMAIN 258 360
FT DOMAIN 135 145
FT DOMAIN 483 772
FT NP BIND 489 497
FT BINDING 519 519
FT ACT_SITE 628 628
FT MOD_RES 659 659
FT DISULFID 65 110
FT DISULFID 181 233
FT DISULFID 280 344
FT CARBOHYD 86 86
FT CARBOHYD 126 126
FT CARBOHYD 148 148
FT CARBOHYD 230 230
FT CARBOHYD 243 243
FT CARBOHYD 267 267
FT CARBOHYD 299 299
FT CARBOHYD 320 320
FT CARBOHYD 333 333
SQ SEQUENCE 823 AA; 92299 MW; 42BF3C4EA02FD43 CRC64;

Query Match 77.6%; Score 83; DB 1; Length 823;
Best Local Similarity 85.7%; Pred. No. 2.9e-06;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 SQAVHKLAKSIPIRQVTVS 22
Db 413 SQAVHKLTKRIPIRQVTVS 433

RESULT 9
FCR2_XENLA
ID FCR2_XENLA STANDARD; PRT; 813 AA.
AC Q03364;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Fibroblast growth factor receptor 2 precursor (EC 2.7.1.112).
GN FGR2-2.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.

```

```

RX MEDLINE=93201992; PubMed=1284237;
RA Friesel R.; Brown S.A.N.;
RT "Spatially restricted expression of fibroblast growth factor
RT Receptor-2 during Xenopus development.";
RL Development 116:1051-1058(1992).
CC -!- FUNCTION: Receptor for acidic and basic fibroblast growth factors.
CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
CC tyrosine phosphate.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- TISSUE SPECIFICITY: EXPRESSED IN THE ANTERIOR NEURAL PLATE
CC IN EARLY NEURULA STAGE EMBRYOS. LATER IN DEVELOPMENT, THE
CC PROTEIN IS ALSO EXPRESSED IN THE EYE ANLAGEN, MIDBRAIN-HINDBRAIN
CC BOUNDARY AND OTIC VESICLE.
CC -!- SIMILARITY: Belongs to the fibroblast growth factor receptor
CC family.
CC -!- SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X65943; CAA46758.1; -
CC PIR; A49123; A49123.
CC HSP; P11362; LFQK.
CC InterPro; IPR007110; Ig-like.
CC InterPro; IPR003598; Ig_c2.
CC InterPro; IPR004074; ILI_receptor/II.
CC InterPro; IPR000719; Prot_kinase.
CC InterPro; IPR001245; Tyr_pkinase.
CC Pfam; PF00047; Ig; 3.
CC Pfam; PF00069; pkinase; 1.
CC PRINTS; PR01536; INTRLNMR12F.
CC PRINTS; PR00109; TYRKINASE.
CC ProDom; PD000001; Prot_kinase; 1.
CC SMART; SM00408; IGC2; 3.
CC SMART; SM00219; TYRK; 1.
CC PROSITE; PS00835; IG LIKE; 3.
CC PROSITE; PS00107; PROTEIN KINASE ATP; 1.
CC PROSITE; PS00011; PROTEIN KINASE DOM; 1.
CC PROSITE; PS00109; PROTEIN KINASE TYR; 1.
KW Receptor; Glycoprotein; Tyrosine-protein kinase; ATP-binding;
KW Transferase; Phosphorylation; Transmembrane; Immunoglobulin domain;
KW Repeat; Signal.
FT SIGNAL 1 14
FT CHAIN 15 813
FT DOMAIN 18 367
FT TRANSMEM 368 388
FT DOMAIN 389 813
FT DOMAIN 35 119
FT DOMAIN 145 237
FT DOMAIN 246 348
FT DOMAIN 471 760
FT NP BIND 477 485
FT BINDING 507 507
FT ACT_SITE 616 616
FT MOD_RES 647 647
FT CARBOHYD 79 79
FT CARBOHYD 115 115
FT CARBOHYD 231 231
FT CARBOHYD 255 255
FT CARBOHYD 287 287
FT CARBOHYD 308 308
FT CARBOHYD 321 321
SQ SEQUENCE 813 AA; 91340 MW; 81543659892A565 CRC64;

Query Match 62.6%; Score 67; DB 1; Length 813;
Best Local Similarity 82.4%; Pred. No. 0.0016;
Matches 14; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

QY 6 VHKLAKSIPIRLQVTVS 22
 DB 405 VHLKTRIPHLRQVTVS 421

RESULT 10
 PGEM_HUMAN STANDARD; PRT; 4391 AA.
 AC P98160; Q16287; Q9H3V5;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Basement membrane-specific heparan sulfate proteoglycan core
 DE protein precursor (HSPG) (perlecan) (PLC).
 GN HSPG2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92112994; PubMed=1730769;
 RA Kallunki P., Tryggvason K.;
 RT "Human basement membrane heparan sulfate proteoglycan core protein: a
 RT 467-kD protein containing multiple domains resembling elements of the
 RT low density lipoprotein receptor; laminin, neural cell adhesion
 RT molecules, and epidermal growth factor.";
 RL J. Cell Biol. 116:559-571 (1992).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX TISSUE=Colon, and Skin;
 RC MEDLINE=92235084; PubMed=1569102;
 RA Murdoch A.D., Dodge G.R., Cohen I., Tuan R.S., Iozzo R.V.;
 RT "Primary structure of the human heparan sulfate proteoglycan from
 RT basement membrane (HSPG2/perlecan): A chimeric molecule with multiple
 RT domains homologous to the low density lipoprotein receptor, laminin,
 RT neural cell adhesion molecules, and epidermal growth factor.";
 RL J. Biol. Chem. 267:8544-8557 (1992).
 RN [3]
 RP SEQUENCE OF 22-4391 FROM N.A., AND VARIANT SJS1 TYR-1532.
 RX MEDLINE=20553141; PubMed=1101850;
 RA Nicole S., Davoine C.-S., Topaloglu H., Cattolico L., Barral D.,
 RA Beighton P., Ben-Hamida C., Hammouda H., Cruaud C., White P.S.,
 RA Sanson D., Urtizberea J.A., Lehmann-Horn F., Weissbach J.,
 RA Hentati F., Fontaine B.;
 RT "Perlecan, the major proteoglycan of basement membranes, is altered in
 RT patients with Schwartz-Jampel syndrome (chondrodystrophic myotonia).";
 RL Nat. Genet. 26:480-483 (2000).
 RN [4]
 RP SEQUENCE OF 1016-1470 FROM N.A.
 RX TISSUE=Colon;
 RC MEDLINE=91365376; PubMed=1679749;
 RA Dodge G.R., Kovalsky I., Chu M.L., Hassell J.R., McBride O.W.,
 RA Yi H.F., Iozzo R.V.;
 RT "Heparan sulfate proteoglycan of human colon: partial molecular
 RT cloning, cellular expression, and mapping of the gene (HSPG2) to the
 RT short arm of human chromosome 1.";
 RL Genomics 10:673-680 (1991).
 RN [5]
 RP SEQUENCE OF 890-1396 FROM N.A.
 RX TISSUE=Fibroblasts;
 RC MEDLINE=92120660; PubMed=168141;
 RA Kallunki P., Eddy R.L., Byers M.G., Kestila M., Shows T.B.,
 RA Tryggvason K.;
 RT "Cloning of human heparan sulfate proteoglycan core protein,
 RT assignment of the gene (HSPG2) to 1p36.1--p35 and identification of
 RT a BamHI restriction fragment length polymorphism.";
 RL Genomics 11:389-396 (1991).
 RN [6]
 RP SEQUENCE OF 1-21 FROM N.A.
 RX MEDLINE=94052171; PubMed=8234307;
 RA Cohen I.R., Graessel S., Murdoch A.D., Iozzo R.V.;

RT "Structural characterization of the complete human perlecan gene and
 RT its promoter.";
 RL Proc. Natl. Acad. Sci. U.S.A. 90:10404-10408 (1993).
 RN [7]
 RP CARBOHYDRATE-LINKAGE SITE ASN-2121.
 RX MEDLINE=22660472; PubMed=12754519;
 RA Zhang H., Li X.-J., Martin D.B., Aebersold R.;
 RT "Identification and quantification of N-linked glycoproteins using
 RT hydrazide chemistry, stable isotope labeling and mass spectrometry.";
 RL Nat. Biotechnol. 21:660-666 (2003).
 CC -!- FUNCTION: This protein is an integral component of basement
 CC membranes. It is responsible for the fixed negative electrostatic
 CC charge and is involved in the charge-selective ultrafiltration
 CC properties. It serves as an attachment substrate for cells.
 CC -!- SUBUNIT: Purified perlecan has a strong tendency to aggregate in
 CC dimers or stellate structures. It interacts with other basement
 CC membrane components such as laminin, prolargin and collagen type
 CC IV.
 CC -!- SUBCELLULAR LOCATION: Extracellular.
 CC -!- TISSUE SPECIFICITY: Found in the basement membranes.
 CC -!- PTM: CONTAINS THREE HEPARAN SULFATE CHAINS AS WELL AS N-LINKED
 CC AND O-LINKED OLIGOSACCHARIDES.
 CC -!- DISEASE: Defects in HSPG2 are the cause of Schwartz-Jampel
 CC syndrome (SJS1) [MIM:255800]; a rare autosomal recessive disorder
 CC characterized by permanent myotonia (prolonged failure of muscle
 CC relaxation) and skeletal dysplasia, resulting in reduced stature,
 CC kyphoscoliosis, bowing of the diaphyses and irregular epiphyses.
 CC -!- SIMILARITY: Contains 4 LDL-receptor class A domains.
 CC -!- SIMILARITY: Contains 11 laminin EGF-like domains.
 CC -!- SIMILARITY: Contains 3 laminin IV domains.
 CC -!- SIMILARITY: Contains 22 immunoglobulin-like C2-type domains.
 CC -!- SIMILARITY: Contains 3 laminin G-like domains.
 CC -!- SIMILARITY: Contains 4 EGF-like domains.
 CC -!- SIMILARITY: Contains 1 SEA domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC EMBL; X62515; CA44373.1; -
 CC EMBL; M85289; AAA52700.1; -
 CC EMBL; AL445785; CAC18534.1; -
 CC EMBL; M64283; AAA52699.1; -
 CC EMBL; S76436; AAB21121.2; -
 CC EMBL; L22078; -; NOT_ANNOTATED_CDS.
 CC PIR; A38096; A38096.
 CC HSP; P00740; IEDM.
 CC Sienna-2DPAGE; P98160; -
 CC Genew; HGNC:5273; HSPG2.
 CC MIM; 142461; -
 CC MIM; 255800; -
 CC InterPro; IPR008985; ConA_like Lec_g1.
 CC InterPro; IPR000742; EGF_2.
 CC InterPro; IPR006209; EGF-like.
 CC InterPro; IPR006210; IEGF.
 CC InterPro; IPR007110; IG-like.
 CC InterPro; IPR003599; IG.
 CC InterPro; IPR003598; IG_c2.
 CC InterPro; IPR003596; IG_v.
 CC InterPro; IPR000034; Laminin_B.
 CC InterPro; IPR002049; Laminin_EGF.
 CC InterPro; IPR001791; Laminin_G.
 CC InterPro; IPR002172; LDL_receptor_A.
 CC InterPro; IPR000082; SEA_domain.
 CC Pfam; PF00008; EGF; 4.
 CC Pfam; PF00047; Ig; 22.
 CC Pfam; PF00052; laminin_B; 3.
 CC Pfam; PF00053; laminin_EGF; 7.
 CC Pfam; PF00054; laminin_G; 3.

RESULT 11

ID AC PGPM MOUSE STANDARD; PRT; 3707 AA.
DT Q05793;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, last sequence update)
DT 10-OCT-2003 (Rel. 42, last annotation update)
DE Basement membrane-specific heparan sulfate proteoglycan core
DE protein precursor (HSPG) (Perlecan) (PLC).
GN HSPG2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
ON NCBI_TaxID=10090;
RX [1]
RX TISSUE=Melanoma;
RX MEDLINE=92078153; PubMed=1744087;
RA Noonan D.M., Fulle A., Valente P., Cai S., Horigan E., Sasaki M.,
RA Yamada Y., Hassell J.R.;
RT "The complete sequence of perlecan, a basement membrane heparan
RT sulfate proteoglycan, reveals extensive similarity with laminin A
RT chain, low density lipoprotein-receptor, and the neural cell adhesion
RT molecule";
RL J. Biol. Chem. 266:22939-22947(1991).
RN [2]
RP SEQUENCE OF 940-1601 AND 1870-2600 FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE=99034110; PubMed=2972708;
RA Noonan D.M., Horigan E.A., Ledbetter S.R., Vogeli G., Sasaki M.,
RA Yamada Y., Hassell J.R.;
RT "Identification of cDNA clones encoding different domains of the
RT basement membrane heparan sulfate proteoglycan";
RL J. Biol. Chem. 263:16379-16387(1988).
CC -!- FUNCTION: This protein is an integral component of basement
CC membranes. It is responsible for the fixed negative electrostatic
CC charge and is involved in the charge-selective ultrafiltration
CC properties. It serves as an attachment substrate for cells.
CC -!- SUBUNIT: Purified perlecan has a strong tendency to aggregate in
CC dimers or stellate structures. It interacts with other basement
CC membrane components such as laminin, prolargin and collagen type
CC IV.
CC -!- SUBCELLULAR LOCATION: Extracellular.
CC -!- TISSUE SPECIFICITY: Found in the basement membranes.
CC -!- PTM: CONTAINS THREE HEPARAN SULFATE CHAINS AS WELL AS N-LINKED
CC AND O-LINKED OLIGOSACCHARIDES.
CC -!- SIMILARITY: Contains 4 LDL-receptor class A domains.
CC -!- SIMILARITY: Contains 11 laminin EGF-like domains.
CC -!- SIMILARITY: Contains 3 laminin IV domains.
CC -!- SIMILARITY: Contains 15 immunoglobulin-like C2-type domains.
CC -!- SIMILARITY: Contains 3 laminin G-like domains.
CC -!- SIMILARITY: Contains 1 EGF-like domain.
CC -!- SIMILARITY: Contains 1 SEA domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M7174; AAA39911.1; -;
CC EMBL; J04054; AAA39899.1; -;
CC EMBL; J04055; AAA39912.1; -;
CC PIR; S18252; S18252.
CC PDB; 1GL4; 28-NOV-01.
CC MGD; MGI196257; Hspg2.
CC GO; GO:0005604; C:basement membrane; IDA.
CC GO; GO:0008104; P:protein localization; IMP.
CC InterPro; IPR008985; ConA like lec_gl.
CC InterPro; IPR000742; EGF 2.
CC InterPro; IPR006209; EGF-like.
CC InterPro; IPR007110; Ig-like.
CC InterPro; IPR003598; Ig_C2.
CC InterPro; IPR000034; Laminin_B.

DR InterPro; IPR002049; Laminin_EGF.
DR InterPro; IPR001791; Laminin_G.
DR InterPro; IPR002172; LDL_receptor_A.
DR InterPro; IPR000082; SEA_domain.
DR Pfam; PF00008; EGF; 4.
DR Pfam; PF00047; Ig; 15.
DR Pfam; PF00052; laminin_B; 3.
DR Pfam; PF00053; laminin_EGF; 7.
DR Pfam; PF00054; laminin_G; 3.
DR Pfam; PF00057; ldl_recept_a; 4.
DR Pfam; PF03390; SEA; 1.
DR PRINTS; PR00261; LDLRECEPTOR.
DR ProDom; PD003031; Laminin_B; 3.
DR SMART; SM00180; EGF_Lam; 7.
DR SMART; SM00408; Igc2; 14.
DR SMART; SM00281; Lamb; 3.
DR SMART; SM00282; LamG; 3.
DR SMART; SM00192; LDLa; 4.
DR SMART; SM00200; SEA; 1.
DR PROSITE; PS00022; EGF 1; 8.
DR PROSITE; PS01186; EGF 2; 5.
DR PROSITE; PS00026; EGF 3; 4.
DR PROSITE; PS00835; IG LIKE; 15.
DR PROSITE; PS00025; LAM G DOMAIN; 3.
DR PROSITE; PS01248; LAMININ_TYPE_EGF; 11.
DR PROSITE; PS01209; LDLRA 1; 4.
DR PROSITE; PS00068; LDLRA 2; 4.
DR PROSITE; PS00024; SEA; 1.
KW Signal; Basement membrane; Proteoglycan; Repeat; Glycoprotein;
KW Heparan sulfate; Laminin EGF-like domain; Immunoglobulin domain;
KW Extracellular matrix; EGF-like domain; 3D-structure.
FT SIGNAL 1 21
FT CHAIN 22 3707
FT DOMAIN 80 194
FT DOMAIN 195 234
FT DOMAIN 281 319
FT DOMAIN 320 359
FT DOMAIN 360 403
FT DOMAIN 404 504
FT DOMAIN 521 530
FT DOMAIN 531 730
FT DOMAIN 731 763
FT DOMAIN 764 813
FT DOMAIN 814 879
FT DOMAIN 879 923
FT DOMAIN 924 933
FT DOMAIN 934 1125
FT DOMAIN 1126 1158
FT DOMAIN 1159 1208
FT DOMAIN 1209 1265
FT DOMAIN 1275 1324
FT DOMAIN 1325 1334
FT DOMAIN 1335 1329
FT DOMAIN 1330 1562
FT DOMAIN 1563 1612
FT DOMAIN 1613 1670
FT DOMAIN 1671 1771
FT DOMAIN 1772 1865
FT DOMAIN 1866 1854
FT DOMAIN 1955 2049
FT DOMAIN 2050 2148
FT DOMAIN 2149 2244
FT DOMAIN 2245 2343
FT DOMAIN 2344 2436
FT DOMAIN 2437 2532
FT DOMAIN 2533 2619
FT DOMAIN 2620 2720
FT DOMAIN 2721 2809
FT DOMAIN 2810 2895
FT DOMAIN 2896 2980
FT DOMAIN 2984 3162
FT DOMAIN 3163 3241
FT DOMAIN 3242 3341
FT DOMAIN 3342 3441
FT DOMAIN 3442 3541
FT DOMAIN 3542 3641
FT DOMAIN 3642 3741
FT DOMAIN 3742 3841
FT DOMAIN 3842 3941
FT DOMAIN 3942 4041
FT DOMAIN 4042 4141
FT DOMAIN 4142 4241
FT DOMAIN 4242 4341
FT DOMAIN 4342 4441
FT DOMAIN 4442 4541
FT DOMAIN 4542 4641
FT DOMAIN 4642 4741
FT DOMAIN 4742 4841
FT DOMAIN 4842 4941
FT DOMAIN 4942 5041
FT DOMAIN 5042 5141
FT DOMAIN 5142 5241
FT DOMAIN 5242 5341
FT DOMAIN 5342 5441
FT DOMAIN 5442 5541
FT DOMAIN 5542 5641
FT DOMAIN 5642 5741
FT DOMAIN 5742 5841
FT DOMAIN 5842 5941
FT DOMAIN 5942 6041
FT DOMAIN 6042 6141
FT DOMAIN 6142 6241
FT DOMAIN 6242 6341
FT DOMAIN 6342 6441
FT DOMAIN 6442 6541
FT DOMAIN 6542 6641
FT DOMAIN 6642 6741
FT DOMAIN 6742 6841
FT DOMAIN 6842 6941
FT DOMAIN 6942 7041
FT DOMAIN 7042 7141
FT DOMAIN 7142 7241
FT DOMAIN 7242 7341
FT DOMAIN 7342 7441
FT DOMAIN 7442 7541
FT DOMAIN 7542 7641
FT DOMAIN 7642 7741
FT DOMAIN 7742 7841
FT DOMAIN 7842 7941
FT DOMAIN 7942 8041
FT DOMAIN 8042 8141
FT DOMAIN 8142 8241
FT DOMAIN 8242 8341
FT DOMAIN 8342 8441
FT DOMAIN 8442 8541
FT DOMAIN 8542 8641
FT DOMAIN 8642 8741
FT DOMAIN 8742 8841
FT DOMAIN 8842 8941
FT DOMAIN 8942 9041
FT DOMAIN 9042 9141
FT DOMAIN 9142 9241
FT DOMAIN 9242 9341
FT DOMAIN 9342 9441
FT DOMAIN 9442 9541
FT DOMAIN 9542 9641
FT DOMAIN 9642 9741
FT DOMAIN 9742 9841
FT DOMAIN 9842 9941
FT DOMAIN 9942 10041
FT DOMAIN 10042 10141
FT DOMAIN 10142 10241
FT DOMAIN 10242 10341
FT DOMAIN 10342 10441
FT DOMAIN 10442 10541
FT DOMAIN 10542 10641
FT DOMAIN 10642 10741
FT DOMAIN 10742 10841
FT DOMAIN 10842 10941
FT DOMAIN 10942 11041
FT DOMAIN 11042 11141
FT DOMAIN 11142 11241
FT DOMAIN 11242 11341
FT DOMAIN 11342 11441
FT DOMAIN 11442 11541
FT DOMAIN 11542 11641
FT DOMAIN 11642 11741
FT DOMAIN 11742 11841
FT DOMAIN 11842 11941
FT DOMAIN 11942 12041
FT DOMAIN 12042 12141
FT DOMAIN 12142 12241
FT DOMAIN 12242 12341
FT DOMAIN 12342 12441
FT DOMAIN 12442 12541
FT DOMAIN 12542 12641
FT DOMAIN 12642 12741
FT DOMAIN 12742 12841
FT DOMAIN 12842 12941
FT DOMAIN 12942 13041
FT DOMAIN 13042 13141
FT DOMAIN 13142 13241
FT DOMAIN 13242 13341
FT DOMAIN 13342 13441
FT DOMAIN 13442 13541
FT DOMAIN 13542 13641
FT DOMAIN 13642 13741
FT DOMAIN 13742 13841
FT DOMAIN 13842 13941
FT DOMAIN 13942 14041
FT DOMAIN 14042 14141
FT DOMAIN 14142 14241
FT DOMAIN 14242 14341
FT DOMAIN 14342 14441
FT DOMAIN 14442 14541
FT DOMAIN 14542 14641
FT DOMAIN 14642 14741
FT DOMAIN 14742 14841
FT DOMAIN 14842 14941
FT DOMAIN 14942 15041
FT DOMAIN 15042 15141
FT DOMAIN 15142 15241
FT DOMAIN 15242 15341
FT DOMAIN 15342 15441
FT DOMAIN 15442 15541
FT DOMAIN 15542 15641
FT DOMAIN 15642 15741
FT DOMAIN 15742 15841
FT DOMAIN 15842 15941
FT DOMAIN 15942 16041
FT DOMAIN 16042 16141
FT DOMAIN 16142 16241
FT DOMAIN 16242 16341
FT DOMAIN 16342 16441
FT DOMAIN 16442 16541
FT DOMAIN 16542 16641
FT DOMAIN 16642 16741
FT DOMAIN 16742 16841
FT DOMAIN 16842 16941
FT DOMAIN 16942 17041
FT DOMAIN 17042 17141
FT DOMAIN 17142 17241
FT DOMAIN 17242 17341
FT DOMAIN 17342 17441
FT DOMAIN 17442 17541
FT DOMAIN 17542 17641
FT DOMAIN 17642 17741
FT DOMAIN 17742 17841
FT DOMAIN 17842 17941
FT DOMAIN 17942 18041
FT DOMAIN 18042 18141
FT DOMAIN 18142 18241
FT DOMAIN 18242 18341
FT DOMAIN 18342 18441
FT DOMAIN 18442 18541
FT DOMAIN 18542 18641
FT DOMAIN 18642 18741
FT DOMAIN 18742 18841
FT DOMAIN 18842 18941
FT DOMAIN 18942 19041
FT DOMAIN 19042 19141
FT DOMAIN 19142 19241
FT DOMAIN 19242 19341
FT DOMAIN 19342 19441
FT DOMAIN 19442 19541
FT DOMAIN 19542 19641
FT DOMAIN 19642 19741
FT DOMAIN 19742 19841
FT DOMAIN 19842 19941
FT DOMAIN 19942 20041
FT DOMAIN 20042 20141
FT DOMAIN 20142 20241
FT DOMAIN 20242 20341
FT DOMAIN 20342 20441
FT DOMAIN 20442 20541
FT DOMAIN 20542 20641
FT DOMAIN 20642 20741
FT DOMAIN 20742 20841
FT DOMAIN 20842 20941
FT DOMAIN 20942 21041
FT DOMAIN 21042 21141
FT DOMAIN 21142 21241
FT DOMAIN 21242 21341
FT DOMAIN 21342 21441
FT DOMAIN 21442 21541
FT DOMAIN 21542 21641
FT DOMAIN 21642 21741
FT DOMAIN 21742 21841
FT DOMAIN 21842 21941
FT DOMAIN 21942 22041
FT DOMAIN 22042 22141
FT DOMAIN 22142 22241
FT DOMAIN 22242 22341
FT DOMAIN 22342 22441
FT DOMAIN 22442 22541
FT DOMAIN 22542 22641
FT DOMAIN 22642 22741
FT DOMAIN 22742 22841
FT DOMAIN 22842 22941
FT DOMAIN 22942 23041
FT DOMAIN 23042 23141
FT DOMAIN 23142 23241
FT DOMAIN 23242 23341
FT DOMAIN 23342 23441
FT DOMAIN 23442 23541
FT DOMAIN 23542 23641
FT DOMAIN 23642 23741
FT DOMAIN 23742 23841
FT DOMAIN 23842 23941
FT DOMAIN 23942 24041
FT DOMAIN 24042 24141
FT DOMAIN 24142 24241
FT DOMAIN 24242 24341
FT DOMAIN 24342 24441
FT DOMAIN 24442 24541
FT DOMAIN 24542 24641
FT DOMAIN 24642 24741
FT DOMAIN 24742 24841
FT DOMAIN 24842 24941
FT DOMAIN 24942 25041
FT DOMAIN 25042 25141
FT DOMAIN 25142 25241
FT DOMAIN 25242 25341
FT DOMAIN 25342 25441
FT DOMAIN 25442 25541
FT DOMAIN 25542 25641
FT DOMAIN 25642 25741
FT DOMAIN 25742 25841
FT DOMAIN 25842 25941
FT DOMAIN 25942 26041
FT DOMAIN 26042 26141
FT DOMAIN 26142 26241
FT DOMAIN 26242 26341
FT DOMAIN 26342 26441
FT DOMAIN 26442 26541
FT DOMAIN 26542 26641
FT DOMAIN 26642 26741
FT DOMAIN 26742 26841
FT DOMAIN 26842 26941
FT DOMAIN 26942 27041
FT DOMAIN 27042 27141
FT DOMAIN 27142 27241
FT DOMAIN 27242 27341
FT DOMAIN 27342 27441
FT DOMAIN 27442 27541
FT DOMAIN 27542 27641
FT DOMAIN 27642 27741
FT DOMAIN 27742 27841
FT DOMAIN 27842 27941
FT DOMAIN 27942 28041
FT DOMAIN 28042 28141
FT DOMAIN 28142 28241
FT DOMAIN 28242 28341
FT DOMAIN 28342 28441
FT DOMAIN 28442 28541
FT DOMAIN 28542 28641
FT DOMAIN 28642 28741
FT DOMAIN 28742 28841
FT DOMAIN 28842 28941
FT DOMAIN 28942 29041
FT DOMAIN 29042 29141
FT DOMAIN 29142 29241
FT DOMAIN 29242 29341
FT DOMAIN 29342 29441
FT DOMAIN 29442 29541
FT DOMAIN 29542 29641
FT DOMAIN 29642 29741
FT DOMAIN 29742 29841
FT DOMAIN 29842 29941
FT DOMAIN 29942 30041
FT DOMAIN 30042 30141
FT DOMAIN 30142 30241
FT DOMAIN 30242 30341
FT DOMAIN 30342 30441
FT DOMAIN 30442 30541
FT DOMAIN 30542 30641
FT DOMAIN 30642 30741
FT DOMAIN 30742 30841
FT DOMAIN 30842 30941
FT DOMAIN 30942 31041
FT DOMAIN 31042 31141
FT DOMAIN 31142 31241
FT DOMAIN 31242 31341
FT DOMAIN 31342 31441
FT DOMAIN 31442 31541
FT DOMAIN 31542 31641
FT DOMAIN 31642 31741
FT DOMAIN 31742 31841
FT DOMAIN 31842 31941
FT DOMAIN 31942 32041
FT DOMAIN 32042 32141
FT DOMAIN 32142 32241
FT DOMAIN 32242 32341
FT DOMAIN 32342 32441
FT DOMAIN 32442 32541
FT DOMAIN 32542 32641
FT DOMAIN 32642 32741
FT DOMAIN 32742 32841
FT DOMAIN 32842 32941
FT DOMAIN 32942 33041
FT DOMAIN 33042 33141
FT DOMAIN 33142 33241
FT DOMAIN 33242 33341
FT DOMAIN 33342 33441
FT DOMAIN 33442 33541
FT DOMAIN 33542 33641
FT DOMAIN 33642 33741
FT DOMAIN 33742 33841
FT DOMAIN 33842 33941
FT DOMAIN 33942 34041
FT DOMAIN 34042 34141
FT DOMAIN 34142 34241
FT DOMAIN 34242 34341
FT DOMAIN 34342 34441
FT DOMAIN 34442 34541
FT DOMAIN 34542 34641
FT DOMAIN 34642 34741
FT DOMAIN 34742 34841
FT DOMAIN 34842 34941
FT DOMAIN 34942 35041
FT DOMAIN 35042 35141
FT DOMAIN 35142 35241
FT DOMAIN 35242 35341
FT DOMAIN 35342 35441
FT DOMAIN 35442 35541
FT DOMAIN 35542 35641
FT DOMAIN 35642 35741
FT DOMAIN 35742 35841
FT DOMAIN 35842 35941
FT DOMAIN 35942 36041
FT DOMAIN 36042 36141
FT DOMAIN 36142 36241
FT DOMAIN 36242 36341
FT DOMAIN 36342 36441
FT DOMAIN 36442 36541
FT DOMAIN 36542 36641
FT DOMAIN 36642 36741
FT DOMAIN 36742 36841
FT DOMAIN 36842 36941
FT DOMAIN 36942 37041
FT DOMAIN 37042 37141
FT DOMAIN 37142 37241
FT DOMAIN 37242 37341
FT DOMAIN 37342 37441
FT DOMAIN 37442 37541
FT DOMAIN 37542 37641
FT DOMAIN 37642 37741
FT DOMAIN 37742 37841
FT DOMAIN 37842 37941
FT DOMAIN 37942 38041
FT DOMAIN 38042 38141
FT DOMAIN 38142 38241
FT DOMAIN 38242 38341
FT DOMAIN 38342 38441
FT DOMAIN 38442 38541
FT DOMAIN 38542 38641
FT DOMAIN 38642 38741
FT DOMAIN 38742 38841
FT DOMAIN 38842 38941
FT DOMAIN 38942 39041
FT DOMAIN 39042 39141
FT DOMAIN 39142 39241
FT DOMAIN 39242 39341
FT DOMAIN 39342 39441
FT DOMAIN 39442 39541
FT DOMAIN 39542 39641
FT DOMAIN 39642 39741
FT DOMAIN 39742 39841
FT DOMAIN 39842 39941
FT DOMAIN 39942 40041
FT DOMAIN 40042 40141
FT DOMAIN 40142 40241
FT DOMAIN 40242 40341
FT DOMAIN 40342 40441
FT DOMAIN 40442 40541
FT DOMAIN 40542 40641
FT DOMAIN 40642 40741
FT DOMAIN 40742 40841
FT DOMAIN 40842 40941
FT DOMAIN 40942 41041
FT DOMAIN 41042 41141
FT DOMAIN 41142 41241
FT DOMAIN 41242 41341
FT DOMAIN 41342 41441
FT DOMAIN 41442 41541
FT DOMAIN 41542 41641
FT DOMAIN 41642 41741
FT DOMAIN 41742 41841
FT DOMAIN 41842 41941
FT DOMAIN 41942 42041
FT DOMAIN 42042 42141
FT DOMAIN 42142 42241
FT DOMAIN 42242 42341
FT DOMAIN 42342 42441
FT DOMAIN 42442 42541
FT DOMAIN 42542 42641
FT DOMAIN 42642 42741
FT DOMAIN 42742 42841
FT DOMAIN 42842 42941
FT DOMAIN 42942 43041
FT DOMAIN 43042 43141
FT DOMAIN 43142 43241
FT DOMAIN 43242 43341
FT DOMAIN 43342 43441
FT DOMAIN 43442 43541
FT DOMAIN 43542 43641
FT DOMAIN 43642 43741
FT DOMAIN 43742 43841
FT DOMAIN 43842 43941
FT DOMAIN 43942 44041
FT DOMAIN 44042 44141
FT DOMAIN 44142 44241
FT DOMAIN 44242 44341
FT DOMAIN 44342 44441
FT DOMAIN 44442 44541
FT DOMAIN 44542 44641
FT DOMAIN 44642 44741
FT DOMAIN 44742 44841
FT DOMAIN 44842 44941
FT DOMAIN 44942 45041
FT DOMAIN 45042 45141
FT DOMAIN 45142 45241
FT DOMAIN 45242 45341
FT DOMAIN 45342 45441
FT DOMAIN 45442 45541
FT DOMAIN 45542 45641
FT DOMAIN 45642 45741
FT DOMAIN 45742 45841
FT DOMAIN 45842 45941
FT DOMAIN 45942 46041
FT DOMAIN 46042 46141
FT DOMAIN 46142 46241
FT DOMAIN 46242 46341
FT DOMAIN 46342 46441
FT DOMAIN 46442 46541
FT DOMAIN 46542 46641
FT DOMAIN 46642 46741
FT DOMAIN 46742 46841
FT DOMAIN 46842 46941
FT DOMAIN 46942 47041
FT DOMAIN 47042 47141
FT DOMAIN 47142 47241
FT DOMAIN 47242 47341
FT DOMAIN 47342 47441
FT DOMAIN 47442 47541
FT DOMAIN 47542 47641
FT DOMAIN 47642 47741
FT DOMAIN 47742 47841
FT DOMAIN 47842 47941
FT DOMAIN 47942 48041
FT DOMAIN 48042 48141
FT DOMAIN 48142 48241
FT DOMAIN 48242 48341
FT DOMAIN 48342 48441
FT DOMAIN 48442 48541
FT DOMAIN 48542 48641
FT DOMAIN 48642 48741
FT DOMAIN 48742 48841
FT DOMAIN 48842 48941
FT DOMAIN 48942 49041
FT DOMAIN 49042 49141
FT DOMAIN 49142 49241
FT DOMAIN 49242 49341
FT DOMAIN 49342 49441
FT DOMAIN 49442 49541
FT DOMAIN 49542 49641
FT DOMAIN 49642 49741
FT DOMAIN 49742 49841
FT DOMAIN 49842 49941
FT DOMAIN 49942 50041
FT DOMAIN 50042 50141
FT DOMAIN 50142 50241
FT DOMAIN 50242 50341
FT DOMAIN 50342 50441
FT DOMAIN 50442 50541
FT DOMAIN 50542 50641
FT DOMAIN 50642 50741
FT DOMAIN 50742 50841
FT DOMAIN 50842 50941
FT DOMAIN 50942 51041
FT DOMAIN 51042 51141
FT DOMAIN 51142 51241
FT DOMAIN 51242 51341
FT DOMAIN 51342 51441
FT DOMAIN 51442 51541
FT DOMAIN 51542 51641
FT DOMAIN 51642 51741
FT DOMAIN 51742 51841
FT DOMAIN 51842 51941
FT DOMAIN 51942 52041
FT DOMAIN 52042 52141
FT DOMAIN 52142 52241
FT DOMAIN 52242 52341
FT DOMAIN 52342 52441
FT DOMAIN 52442 52541
FT DOMAIN 52542 52641
FT DOMAIN 52642 52741
FT DOMAIN 52742 52841
FT DOMAIN 52842 52941
FT DOMAIN 52942 53041
FT DOMAIN 53042 53141
FT DOMAIN 53142 53241
FT DOMAIN 53242 53341
FT DOMAIN 53342 53441
FT DOMAIN 53442 53541
FT DOMAIN 53542 53641
FT DOMAIN 53642 53741
FT DOMAIN 53742 53841
FT DOMAIN 53842 53941
FT DOMAIN 53942 54041
FT DOMAIN 54042 54141
FT DOMAIN 54142 54241
FT DOMAIN 54242 54341
FT DOMAIN 54342 54441
FT DOMAIN 54442 54541
FT DOMAIN 54542 54641
FT DOMAIN 54642 54741
FT DOMAIN 54742 54841
FT DOMAIN 54842 54941
FT DOMAIN 54942 55041
FT DOMAIN 55042 55141
FT DOMAIN 55142 55241
FT DOMAIN 55242 55341
FT DOMAIN 55342 55441
FT DOMAIN 55442 55541
FT DOMAIN 55542 55641
FT DOMAIN 55642 55741
FT DOMAIN 55742 55841
FT DOMAIN 55842 55941
FT DOMAIN 55942 56041
FT DOMAIN 56042 56141
FT DOMAIN 56142 56241
FT DOMAIN 56242 56341
FT DOMAIN 56342 56441
FT DOMAIN 56442 56541
FT DOMAIN 56542 56641
FT DOMAIN 56642 56741
FT DOMAIN 56742 56841
FT DOMAIN 56842 56941
FT DOMAIN 56942 57041
FT DOMAIN 57042 57141
FT DOMAIN 57142 57241
FT DOMAIN 57242 57341
FT DOMAIN 57342 57441
FT DOMAIN 57442 57541
FT DOMAIN 57542 57641
FT DOMAIN 57642 57741
FT DOMAIN 57742 57841
FT DOMAIN 57842 57941
FT DOMAIN 57942 58041
FT DOMAIN 58042 58141
FT DOMAIN 58142 58241
FT DOMAIN 58242 58341
FT DOMAIN 58342 58441
FT DOMAIN 58442 58541
FT DOMAIN 58542 58641
FT DOMAIN 58642 58741
FT DOMAIN 58742 58841
FT DOMAIN 58842 58941
FT DOMAIN 58942 59041
FT DOMAIN 59042 59141
FT DOMAIN 59142 59241
FT DOMAIN 59242 59341
FT DOMAIN 59342 59441
FT DOMAIN 59442 59541
FT DOMAIN 59542 59641
FT DOMAIN 59642 59741
FT DOMAIN 59742 59841
FT DOMAIN 59842 59941
FT DOMAIN 59942 60041
FT DOMAIN 60042 60141
FT DOMAIN 60142 60241
FT DOMAIN 60242 60341
FT DOMAIN 60342 60441
FT DOMAIN 60442 60541
FT DOMAIN 60542 60641
FT DOMAIN 60642 60741
FT DOMAIN 60742 60841
FT DOMAIN 60842 60941
FT DOMAIN 60942 61041
FT DOMAIN 61042 61141
FT DOMAIN 61142 61241
FT DOMAIN 61242 61341
FT DOMAIN 61342 61441
FT DOMAIN 61442 61541
FT DOMAIN 61542 61641
FT DOMAIN 61642 61741
FT DOMAIN 61742 61841
FT DOMAIN 61842 61941
FT DOMAIN 61942 62041
FT DOMAIN 62042 62141
FT DOMAIN 62142 62241
FT DOMAIN 62242 62341
FT DOMAIN 62342 62441
FT DOMAIN 62442 62541
FT DOMAIN 62542 62641
FT DOMAIN 62642 62741
FT DOMAIN 62742 62841
FT DOMAIN 62842 62941
FT DOMAIN 62942 63041
FT DOMAIN 63042 63141
FT DOMAIN 63142 63241
FT DOMAIN 63242 63341
FT DOMAIN 63342 63441
FT DOMAIN 63442 63541
FT DOMAIN 63542 63641
FT DOMAIN 63642 63741
FT DOMAIN 63742 63841
FT DOMAIN 63842 63941
FT DOMAIN 63942 64041
FT DOMAIN 64042 64141
FT DOMAIN 64142 64241
FT DOMAIN 64242 64341
FT DOMAIN 64342 64441
FT DOMAIN 64442 64541
FT DOMAIN 645

```
FT DOMAIN 3245 3425 LAMININ G-LIKE 2.
FT DOMAIN 3518 3705 LAMININ G-LIKE 3.
FT SITE 65 67 HEPARAN SULFATE (POTENTIAL).
FT SITE 71 73 HEPARAN SULFATE (POTENTIAL).
FT SITE 76 78 HEPARAN SULFATE (POTENTIAL).
FT SITE 3615 3617 MEDIATES MOTOR NEURON ATTACHMENT
(POTENTIAL).
FT DISULFID 199 212 BY SIMILARITY.
FT DISULFID 206 225 BY SIMILARITY.
FT DISULFID 219 234 BY SIMILARITY.
FT DISULFID 285 297 BY SIMILARITY.
FT DISULFID 292 310 BY SIMILARITY.
FT DISULFID 304 319 BY SIMILARITY.
FT DISULFID 325 337 BY SIMILARITY.
FT DISULFID 332 350 BY SIMILARITY.
FT DISULFID 344 359 BY SIMILARITY.
FT DISULFID 368 381 BY SIMILARITY.
FT DISULFID 375 394 BY SIMILARITY.
FT DISULFID 388 403 BY SIMILARITY.
FT DISULFID 428 479 BY SIMILARITY.
FT DISULFID 764 773 BY SIMILARITY.
FT DISULFID 766 780 BY SIMILARITY.
FT DISULFID 792 792 BY SIMILARITY.
FT DISULFID 795 811 BY SIMILARITY.
FT DISULFID 814 829 BY SIMILARITY.
FT DISULFID 816 839 BY SIMILARITY.
FT DISULFID 842 851 BY SIMILARITY.
FT DISULFID 854 869 BY SIMILARITY.
FT DISULFID 1159 1168 BY SIMILARITY.
FT DISULFID 1161 1175 BY SIMILARITY.
FT DISULFID 1178 1187 BY SIMILARITY.
FT DISULFID 1190 1206 BY SIMILARITY.
FT DISULFID 1209 1224 BY SIMILARITY.
FT DISULFID 1211 1234 BY SIMILARITY.
FT DISULFID 1237 1246 BY SIMILARITY.
FT DISULFID 1249 1263 BY SIMILARITY.
FT DISULFID 1275 1287 BY SIMILARITY.
FT DISULFID 1277 1293 BY SIMILARITY.
FT DISULFID 1295 1304 BY SIMILARITY.
FT DISULFID 1307 1322 BY SIMILARITY.
FT DISULFID 1563 1572 BY SIMILARITY.
FT DISULFID 1565 1579 BY SIMILARITY.
FT DISULFID 1582 1591 BY SIMILARITY.
FT DISULFID 1594 1610 BY SIMILARITY.
FT DISULFID 1613 1628 BY SIMILARITY.
FT DISULFID 1615 1638 BY SIMILARITY.
FT DISULFID 1641 1650 BY SIMILARITY.
FT DISULFID 1653 1668 BY SIMILARITY.
FT DISULFID 1792 1839 BY SIMILARITY.
FT DISULFID 1886 1932 BY SIMILARITY.
FT DISULFID 1976 2021 BY SIMILARITY.
FT DISULFID 2073 2118 BY SIMILARITY.
FT DISULFID 2170 2215 BY SIMILARITY.
FT DISULFID 2268 2313 BY SIMILARITY.

Query Match 42.1%; Score 45; DB 1; Length 3707;
Best Local Similarity 44.4%; Pred. No. 53;
Matches 8; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 1 H5QMAVHKLAKSIPLRRQ 18
DB 2275 HAQVTHKRGSSLPARHQ 2292

RESULT 13
VAL2 TYLCV STANDARD; PRT; 135 AA.
AC P27252;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE A12 protein (C2 protein).
GN C2.

Tomato yellow leaf curl virus (TYLCV).
Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
NCBI_TaxID=10832;
[1]
SEQUENCE FROM N.A.
MEDLINE=92024070; PubMed=1926771;
RA Navot N., Fichersky E., Zeidan M., Zamir D., Czosnek H.;
RT "Tomato yellow leaf curl virus: a whitefly-transmitted geminivirus
with a single genomic component.";
RL Virology 185:151-161(1991).
CC -!- SIMILARITY: Belongs to the geminiviruses AL2 protein family.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch)
-----
CC EMBL; X15656; CAA33689.1; -.
DR PIR; C40779; QCCVC4.
DR InterPro; IPR000942; Gemini AL2.
DR Pfam; PF01440; Gemini AL2; 1.
DR PRINTS; PR00230; GEMCOATAL2.
DR ProDom; PD001117; Gemini AL2; 1.
SQ SEQUENCE 135 AA; 15611 MW; F111C8C2P7E9DD32 CRC64;

Query Match 40.7%; Score 43.5; DB 1; Length 135;
Best Local Similarity 45.0%; Pred. No. 2.6;
Matches 9; Conservative 6; Mismatches 2; Indels 3; Gaps 1;

QY 2 SOMAV--HKLAKSIPLRRQ 18
DB 12 SQVSIKVQHKIAKKPIRRK 31

RESULT 14
AROB BACCR
ID AROB BACCR STANDARD; PRT; 361 AA.
AC Q81FQ2;
DT 15-MAR-2004 (Rel. 43, Created)
DT 15-MAR-2004 (Rel. 43, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE 3-dehydroquininate synthase (EC 4.2.3.4).
GN AROB OR BC1517.
OS Bacillus cereus (strain ATCC 14579 / DSM 31).
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=226900;
RN [1]
SEQUENCE FROM N.A.
MEDLINE=22608415; PubMed=12721630;
RA Ivanova N., Sorokin A., Anderson I., Galleron N., Candelon B.,
Kapatral V., Bhattacharya A., Reznik G., Mikhailova N., Lapidus A.,
Chu L., Mazur M., Goltsman E., Larsen N., D'Souza M., Walunas T.,
Grechkin Y., Fusch G., Haseikorn R., Fongstein M., Ehrlich S.D.,
RA Overbeek R., Kyrpides N.;
RT "Genome sequence of Bacillus cereus and comparative analysis with
Bacillus anthracis.";
RL Nature 423:87-91(2003).
CC -!- CATALYTIC ACTIVITY: 3-deoxy-arabino-heptulosonate 7-phosphate = 3-
dehydroquininate + phosphate.
CC -!- COFACTOR: NAD and a divalent metal cation (By similarity).
CC -!- PATHWAY: Aromatic amino acids biosynthesis; shikimate pathway;
second step.
CC -!- SUBUNIT: Monomer (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC -!- SIMILARITY: Belongs to the dehydroquininate synthase family.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
```

CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

DR EMBL; AE017002; AAP08497.1; -
DR HAVAP; MF_00110; -; 1.
DR InterPro; IPR002658; DHQ synthase.
DR Pfam; PF01761; DHQ synthase; 1.
DR TIGRFAMs; TIGR01357; aroB; 1.
KW Aromatic amino acid biosynthesis; Lyase; NAD; Complete proteome.
SQ SEQUENCE 361 AA; 39974 MW; E456B4259AA5ED1 CRC64;

Query Match 40.2%; Score 43; DB 1; Length 361;
Best Local Similarity 36.8%; Pred.No. 9.4;
Matches 7; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

QY 1 HSMQAVHKLAKSIPLRRQV 19
DB 211 HDEKLIHLTKAIPVKANV 229

RESULT 15

EH22 HUMAN STANDARD; PRT; 543 AA.
AC Q9NZN4;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE EH-domain containing protein 2.
GN EHD2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=20139435; PubMed=10673336;
RA Pohl U., Smith J.S., Tachibana I., Ueki K., Lee H.K., Ramaswamy S.,
RA Wu Q., Mohrenweiser H.W., Jenkins R.B., Louis D.N.;
RT "EHD2, EHD3, and EHD4 encode novel members of a highly conserved
RT family of EH domain-containing proteins.";
RL Genomics 63:255-262(2000).
CC -!- TISSUE SPECIFICITY: Highly expressed in heart and moderately
CC expressed in placenta, lung, and skeletal muscle.
CC -!- SIMILARITY: Contains 1 EH domain.
CC -!- SIMILARITY: Contains 1 EF-hand calcium-binding domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

DR EMBL; AF181263; AAF40470.1; -
DR Genew; HGNC:3243; EHD2.
DR MIM; 605890; -
DR GO; GO:0005634; C:nucleus; TAS.
DR GO; GO:0003676; F:nucleic acid binding; TAS.
DR InterPro; IPR002048; EF-hand.
DR InterPro; IPR00261; EPS15_homology.
DR SMART; SM00027; EH; 1.
DR PROSITE; PS00018; EF_HAND; 1.
DR PROSITE; PS00031; EH; 1.
KW Calcium-binding; ATP-binding.
FT NP_BIND 65 72 ATP (POTENTIAL).
FT DOMAIN 449 537 EH.
FT CA_BIND 494 505 EF-HAND (POTENTIAL).
SQ SEQUENCE 543 AA; 61106 MW; 7A95ADAB9C7A76CC CRC64;

Query Match 39.3%; Score 42; DB 1; Length 543;

Best Local Similarity 50.0%; Pred.No. 22;
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;
QY 3 QMAVHKLAKSIPLRRQ 18
DB 384 EMLTHDIAKLMPLLRQ 399

Search completed: August 2, 2004, 09:29:57
Job time : 3.31579 secs

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143	144	145	146	147	148	149	150	151	152	153	154	155	156	157	158	159	160	161	162	163	164	165	166	167	168	169	170	171	172	173	174	175	176	177	178	179	180	181	182	183	184	185	186	187	188	189	190	191	192	193	194	195	196	197	198	199	200	201	202	203	204	205	206	207	208	209	210	211	212	213	214	215	216	217	218	219	220	221	222	223	224	225	226	227	228	229	230	231	232	233	234	235	236	237	238	239	240	241	242	243	244	245	246	247	248	249	250	251	252	253	254	255	256	257	258	259	260	261	262	263	264	265	266	267	268	269	270	271	272	273	274	275	276	277	278	279	280	281	282	283	284	285	286	287	288	289	290	291	292	293	294	295	296	297	298	299	300	301	302	303	304	305	306	307	308	309	310	311	312	313	314	315	316	317	318	319	320	321	322	323	324	325	326	327	328	329	330	331	332	333	334	335	336	337	338	339	340	341	342	343	344	345	346	347	348	349	350	351	352	353	354	355	356	357	358	359	360	361	362	363	364	365	366	367	368	369	370	371	372	373	374	375	376	377	378	379	380	381	382	383	384	385	386	387	388	389	390	391	392	393	394	395	396	397	398	399	400	401	402	403	404	405	406	407	408	409	410	411	412	413	414	415	416	417	418	419	420	421	422	423	424	425	426	427	428	429	430	431	432	433	434	435	436	437	438	439	440	441	442	443	444	445	446	447	448	449	450	451	452	453	454	455	456	457	458	459	460	461	462	463	464	465	466	467	468	469	470	471	472	473	474	475	476	477	478	479	480	481	482	483	484	485	486	487	488	489	490	491	492	493	494	495	496	497	498	499	500	501	502	503	504	505	506	507	508	509	510	511	512	513	514	515	516	517	518	519	520	521	522	523	52
--	---	---	---	---	---	---	---	---	---	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	----

Query Match 100.0%; Score 107; DB 11; Length 203;
Best Local Similarity 100.0%; Pred. No. 2.6e-10;

```
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HSQMAVHKLAKSIPLRRQVTVS 22
DB 76 HSQMAVHKLAKSIPLRRQVTVS 97

RESULT 2
ID Q02063 PRELIMINARY; PRT; 662 AA.
AC Q02063;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Heparin-binding growth factor receptor.
GN HBGF-R.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=91126480; PubMed=1846977;
RA Hou J., Kan M., McKeehan K., McBride G., Adams P., McKeehan W.L.;
RT "Fibroblast growth factor receptors from liver vary in three
structural domains.";
RL Science 251:665-668(1991).
DR EMBL; M63888; AAA35959.1; -.
DR PIR; C40862; C40862.
DR HSSP; P11362; 1FGK.
DR GO; GO:0016021; C:integral to membrane; NAS.
DR GO; GO:0005524; F:ATP binding; NAS.
DR GO; GO:0005007; F:fibroblast growth factor receptor activity; NAS.
DR GO; GO:0008201; F:heparin binding; NAS.
DR GO; GO:0004713; F:protein-tyrosine kinase activity; NAS.
DR GO; GO:0006468; P:protein amino acid phosphorylation; NAS.
DR InterPro; IPR007110; IG_LIKE.
DR InterPro; IPR003598; IG_C2.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00408; IGC2; 3.
DR PROSITE; PS00835; IG_LIKE; 3.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR ATP-binding; Immunoglobulin domain; Receptor; Transferase.
SQ SEQUENCE 662 AA; 73462 MW; E90509648FF937D9 CRC64;

Query Match 100.0%; Score 107; DB 4; Length 662;
Best Local Similarity 100.0%; Pred. No. 8,6e-10;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HSQMAVHKLAKSIPLRRQVTVS 22
DB 409 HSQMAVHKLAKSIPLRRQVTVS 430

RESULT 3
ID Q8CFK8 PRELIMINARY; PRT; 731 AA.
AC Q8CFK8;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Similar to fibroblast growth factor receptor 1 (fms-related tyrosine
kinase 2, Pfeiffer syndrome).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Balb/C;
RX MEDLINE=20283643; PubMed=10821861;
RA Beer H.-D., Vindevoghel L., Gait M.J., Revest J.-M., Duan D.R.,
RA Mason I., Dickson C., Werner S.;
RT "Fibroblast growth factor (FGF) receptor 1-iiib is a naturally
occurring functional receptor for FGFs that is preferentially expressed
in the skin and the brain.";
RL J. Biol. Chem. 275:16091-16097(2000).
DR EMBL; AF176552; AAF05312.1; -.
DR PIR; B42057; B42057.
DR HSSP; P11362; 1FGK.
DR MGD; MGI:95522; Fcfr1.
DR GO; GO:0007420; P:brain development; IMP.
DR GO; GO:004472; P:primer ear morphogenesis; IMP.
DR GO; GO:0007435; P:salivary gland morphogenesis; IMP.
DR InterPro; IPR007110; IG_LIKE.
DR InterPro; IPR003598; IG_C2.
```

```
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Breast tumor;
RA Strausberg R.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC033447; AAH33447.1; -.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR003599; IG.
DR InterPro; IPR007110; IG_LIKE.
DR InterPro; IPR003598; IG_C2.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_kinase.
DR InterPro; IPR001245; Tyr_pkinase.
DR InterPro; IPR008266; Tyr_pkinase_AS.
DR Pfam; PF00047; IG; 2.
DR Pfam; PF00069; pkinase; 1.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00409; IG; 2.
DR SMART; SM00408; IGC2; 2.
DR SMART; SM00220; S_TK; 1.
DR SMART; SM00219; TyRK; 1.
DR PROSITE; PS00835; IG_LIKE; 2.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
KW Receptor.
SQ SEQUENCE 731 AA; 81898 MW; 2E88793289A97818 CRC64;

Query Match 100.0%; Score 107; DB 11; Length 731;
Best Local Similarity 100.0%; Pred. No. 9,5e-10;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HSQMAVHKLAKSIPLRRQVTVS 22
DB 318 HSQMAVHKLAKSIPLRRQVTVS 339

RESULT 4
Q0QZM7
ID Q0QZM7 PRELIMINARY; PRT; 733 AA.
AC Q0QZM7;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Fibroblast growth factor receptor 1-iiib.
GN FGFR1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Balb/C;
RX MEDLINE=20283643; PubMed=10821861;
RA Beer H.-D., Vindevoghel L., Gait M.J., Revest J.-M., Duan D.R.,
RA Mason I., Dickson C., Werner S.;
RT "Fibroblast growth factor (FGF) receptor 1-iiib is a naturally
occurring functional receptor for FGFs that is preferentially expressed
in the skin and the brain.";
RL J. Biol. Chem. 275:16091-16097(2000).
DR EMBL; AF176552; AAF05312.1; -.
DR PIR; B42057; B42057.
DR HSSP; P11362; 1FGK.
DR MGD; MGI:95522; Fcfr1.
DR GO; GO:0007420; P:brain development; IMP.
DR GO; GO:004472; P:primer ear morphogenesis; IMP.
DR GO; GO:0007435; P:salivary gland morphogenesis; IMP.
DR InterPro; IPR007110; IG_LIKE.
DR InterPro; IPR003598; IG_C2.
```



```
DR InterPro: IPR000719; Prot kinase.
DR InterPro: IPR001245; Tyr_pkinase.
DR InterPro: IPR008266; Tyr_pkinase_AS.
DR Pfam: PF00047; Ig; 2.
DR Pfam: PF00089; pkinase; 1.
DR PRINTS: PR00109; TYRKINASE.
DR ProDom: PD000001; Prot kinase; 1.
DR SMART: SM00408; IGC2; 2.
DR SMART: SM00219; TyKc; 1.
DR PROSITE: PS00835; IG LIKE; 2.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: PS00011; PROTEIN_KINASE_DOM; 1.
DR PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
DR ATP-binding; Immunoglobulin domain; Kinase; Receptor; Transferase;
KW Tyrosine-protein kinase.
SW Tyrosine-protein kinase.
SQ SEQUENCE 733 AA; 82122 MW; D57EB7642678D293 CRC64;

Query Match 100.0%; Score 107; DB 11; Length 733;
Best Local Similarity 100.0%; Pred. No. 9.6e-10;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HSQNAVHKLAKSIPLRRQVTVS 22
DB 320 HSQNAVHKLAKSIPLRRQVTVS 341

RESULT 5
Q60830 PRELIMINARY; PRT; 733 AA.
AC Q60830;
DT 01-NOV-1996 (TREMELrel. 01, Created)
DT 01-NOV-1996 (TREMELrel. 01, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE Fibroblast growth factor receptor-1, short isoform.
GN FGFR1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Swiss Webster; TISSUE=Heart;
RX MEDLINE=95205422; PubMed=7897669;
RA Jin Y.; Pasumathri K.B.; Bock M.E.; Lytras A.; Kardami E.;
RA Cattini P.A.;
RT "Cloning and expression of fibroblast growth factor receptor-1
RT isoforms in the mouse heart: evidence for isoform switching during
RT heart development."
RL J. Mol. Cell. Cardiol. 26:1449-1459(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Swiss Webster; TISSUE=Heart;
RA Jin Y.;
RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL: U23445; AAC52183.1; -.
DR PIR: I49293; I49293.
DR HSP: P11362; IFGK.
DR MGD: MGI:95522; Fgfr1.
DR GO: GO:0007420; P:brain development; IMP.
DR GO: GO:0042472; P:inner ear morphogenesis; IMP.
DR GO: GO:0007435; P:salivary gland morphogenesis; IMP.
DR InterPro: IPR007110; Ig-like.
DR InterPro: IPR003598; Ig_c2.
DR InterPro: IPR000719; Prot kinase.
DR InterPro: IPR001245; Tyr_pkinase.
DR InterPro: IPR008266; Tyr_pkinase_AS.
DR Pfam: PF00047; Ig; 2.
DR Pfam: PF00069; pkinase; 1.
DR PRINTS: PR00109; TYRKINASE.
DR ProDom: PD000001; Prot kinase; 1.
DR SMART: SM00408; IGC2; 2.
DR SMART: SM00219; TyKc; 1.
DR PROSITE: PS00835; IG LIKE; 2.

DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: PS00011; PROTEIN_KINASE_DOM; 1.
DR PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
DR ATP-binding; Immunoglobulin domain; Kinase; Receptor; Transferase;
KW Tyrosine-protein kinase.
SW Tyrosine-protein kinase.
SQ SEQUENCE 733 AA; 82168 MW; 23F35C024C3DE7B6 CRC64;

Query Match 100.0%; Score 107; DB 11; Length 733;
Best Local Similarity 100.0%; Pred. No. 9.6e-10;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HSQNAVHKLAKSIPLRRQVTVS 22
DB 320 HSQNAVHKLAKSIPLRRQVTVS 341

RESULT 6
Q80T10 PRELIMINARY; PRT; 733 AA.
AC Q80T10;
DT 01-JUN-2003 (TREMELrel. 24, Created)
DT 01-JUN-2003 (TREMELrel. 24, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE 18 days pregnant adult female placenta and extra embryonic tissue
DE cDNA, RIKEN full-length enriched library, clone:3830408H21
DE product: fibroblast growth factor receptor 1, full insert
DE sequence.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Placenta, and Extraembryonic tissue;
RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
RA Fukuda S., Furuno M., Hanegaki T., Hara A., Hashizume M.,
RA Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,
RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
RA Kato H., Kawai J., Kojima T., Miyazaki A., Kondo S., Kondo M., Koya S.,
RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
RA Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,
RA Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,
RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Placenta, and Extraembryonic tissue;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs."
RL Nature 420:563-573(2002).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Placenta, and Extraembryonic tissue;
RA RIKEN FANTOM Consortium,
RT "Functional annotation of a full-length mouse cDNA collection."
RL Nature 409:685-690(2001).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Placenta, and Extraembryonic tissue;
RX MEDLINE=99279253; PubMed=10349636;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning."
RL Meth. Enzymol. 303:19-44(1999).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Placenta, and Extraembryonic tissue;
RX MEDLINE=20499374; PubMed=11042159;
```

RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Kono H., Okazaki Y., Muramatsu M., Hayashizaki Y.,
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes."
RL Genome Res. 10:1617-1630(2000).
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=CS7BL/6J; TISSUE=Placenta, and Extraembryonic tissue;
RX MEDLINE=20530913; PubMed=11078661;
RA Shibata K., Itoh M., Aizawa K., Nagao S., Sasaki N., Carninci P.,
RA Kono H., Akiyama J., Nishi K., Kitsuai T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura K., Hazama M., Nishine T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watabiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuda S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.,
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer."
RL Genome Res. 10:1757-1771(2000).
DR EMBL; AK028354; BAC25899.1; -;
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR003599; IG.
DR InterPro; IPR007110; IG-like.
DR InterPro; IPR003598; IG_c2.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_kinase.
DR InterPro; IPR001245; Tyr_pkinase.
DR InterPro; IPR008266; Tyr_pkinase_AS.
DR Pfam; PF00047; ig; 3.
DR Pfam; PF00069; pkinase; 1.
DR PRINTS; PR00109; TYRKINASE.
DR PRODOM; PD000001; Prot_kinase; 1.
DR SMART; SM00409; IG; 3.
DR SMART; SM00408; IGC2; 3.
DR SMART; SM00220; S_TKC; 1.
DR SMART; SM00219; Tyrc; 1.
DR PROSITE; PS00835; IG_LIKE; 3.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR ATP-binding; Immunoglobulin domain; Receptor; Transferase.
SQ SEQUENCE 820 AA; 91451 MW; 314F5798B05CC95C CRC64;
Query Match 100.0%; Score 107; DB 4; Length 820;
Best Local Similarity 100.0%; Pred. No. 1.1e-09; Indels 0; Gaps 0;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 HSQMAVHKLAKSIPLRQVTVS 22
DB 407 HSQMAVHKLAKSIPLRQVTVS 428
|||||
RESULT 8
Q8CIN9 PRELIMINARY; PRT; 820 AA.
AC Q8CIN9;
DT 01-VAR-2003 (TrEMBLrel. 23, Created)
DT 01-VAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Similar to fibroblast growth factor receptor 1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]_TaxID=10090;
RP SEQUENCE FROM N.A.
RC STRAIN=FVB/N;
RA Strausberg R.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC010200; AAH10200.1; -;
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR003599; IG.
DR InterPro; IPR007110; IG-like.
DR InterPro; IPR003598; IG_c2.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_kinase.
DR InterPro; IPR001245; Tyr_pkinase.
DR InterPro; IPR008266; Tyr_pkinase_AS.
DR Pfam; PF00047; ig; 3.
DR Pfam; PF00069; pkinase; 1.
DR PRINTS; PR00109; TYRKINASE.
DR PRODOM; PD000001; Prot_kinase; 1.
DR SMART; SM00409; IG; 3.
DR SMART; SM00408; IGC2; 3.
QY 1 HSQMAVHKLAKSIPLRQVTVS 22
DB 320 HSQMAVHKLAKSIPLRQVTVS 341
|||||
RESULT 7
Q8N685 PRELIMINARY; PRT; 820 AA.
AC Q8N685;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Similar to fibroblast growth factor receptor 1 (fms-related tyrosine
DE kinase 2; Pfeiffer syndrome).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Pancreas;
RA Strausberg R.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC018128; AAH18128.1; -;

DR SMART: SM00220; S_TKC; 1.
 DR SMART: SM00219; TYR_KC; 1.
 DR PROSITE; PS00835; IG_LIKE; 3.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
 DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
 KW Receptor.
 SQ SEQUENCE 820 AA; 91668 MW; 58319BD3BEA9D34 CRC64;

Query Match 100.0%; Score 107; DB 11; Length 820;
 Best Local Similarity 100.0%; Pred. No. 1.1e-09;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HSQMAVHKLAKSIPLRRQVTVS 22
 DB 407 HSQMAVHKLAKSIPLRRQVTVS 428

RESULT 9
 Q60818 PRELIMINARY; PRT; 822 AA.
 AC Q60818, 1996 (TREMELrel. 01, Created)
 DT 01-NOV-1996 (TREMELrel. 01, Last sequence update)
 DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
 DE Fibroblast growth factor receptor-1, long isoform.
 GN FGFR1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Swiss Webster; TISSUE=Heart;
 RX MEDLINE=95205422; Pubmed=7897669;
 RA Jin Y.; Pasumamathi K.B.; Bock M.B.; Lytras A.; Kardami E.,
 Cattini P.A.;
 RT "Cloning and expression of fibroblast growth factor receptor-1
 RT isoforms in the mouse heart: evidence for isoform switching during
 RT heart development.";
 RL J. Mol. Cell. Cardiol. 26:1449-1459(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Swiss Webster; TISSUE=Heart;
 RA Jin Y.;
 RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.
 DR EMBL; U22324; AAC52182.1; -.
 DR PIR; I49289; I49289.
 DR HSP; P11362; IFGK.
 DR MG; MG195522; Fgfr1.
 DR GO; GO:0007420; P:brain development; IMP.
 DR GO; GO:0042472; P:inner ear morphogenesis; IMP.
 DR GO; GO:0007435; P:salivary gland morphogenesis; IMP.
 DR InterPro; IPR007110; IG-like.
 DR InterPro; IPR003598; IG_c2.
 DR InterPro; IPR00719; Prot_kinase.
 DR InterPro; IPR001245; Tyr_pkinase.
 DR InterPro; IPR008266; Tyr_pkinase_AS.
 DR Pfam; PF00047; ig; 3.
 DR PRINTS; PR00109; TYRKINASE.
 DR ProDom; PD000001; Prot_kinase; 1.
 DR SMART; SMC0408; IGC2; 3.
 DR SMART; SMC0219; TYRK; 1.
 DR PROSITE; PS00835; IG_LIKE; 3.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
 DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
 KW ATP-binding; Immunoglobulin domain; Kinase; Receptor; Transferase;
 TYROSINE-protein kinase.
 FT CHAIN 21 822 FIBROBLAST GROWTH FACTOR RECEPTOR-1, LONG
 ISOFORM.
 SQ SEQUENCE 822 AA; 91938 MW; A6BEGABBAED450F5 CRC64;

Query Match 100.0%; Score 107; DB 11; Length 822;
 Best Local Similarity 100.0%; Pred. No. 1.1e-09;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HSQMAVHKLAKSIPLRRQVTVS 22
 DB 409 HSQMAVHKLAKSIPLRRQVTVS 430

RESULT 10
 Q03836 PRELIMINARY; PRT; 724 AA.
 AC Q03836;
 DT 01-NOV-1996 (TREMELrel. 01, Created)
 DT 01-NOV-1996 (TREMELrel. 01, Last sequence update)
 DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
 DE Fibroblast growth factor receptor precursor.
 GN XFGPRA2.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
 OC Xenopodinae; Xenopus.
 OX NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91203867; Pubmed=1850097;
 RA Friesel R.E.; David I.B.;
 RT "cDNA cloning and developmental expression of fibroblast growth factor
 RT receptors from Xenopus laevis.";
 RL Mol. Cell. Biol. 11:2481-2488(1991).
 DR EMBL; M62322; AAA49993.1; -.
 DR HSP; P11362; IFGK.
 DR GO; GO:0005524; P:ATP binding; IEA.
 DR GO; GO:0004713; P:protein-tyrosine kinase activity; IEA.
 DR GO; GO:0004872; P:receptor activity; IEA.
 DR GO; GO:0016740; P:transferase activity; IEA.
 DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
 DR InterPro; IPR007110; IG-like.
 DR InterPro; IPR003598; IG_c2.
 DR InterPro; IPR000719; Prot_kinase.
 DR InterPro; IPR001245; Tyr_pkinase.
 DR InterPro; IPR008266; Tyr_pkinase_AS.
 DR Pfam; PF00047; ig; 2.
 DR Pfam; PF00069; pkinase; 1.
 DR ProDom; PD000001; Prot_kinase; 1.
 DR SMART; SMC0408; IGC2; 2.
 DR SMART; SMC00219; TYRK; 1.
 DR PROSITE; PS00835; IG_LIKE; 2.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
 DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
 KW ATP-binding; Immunoglobulin domain; Kinase; Receptor; Signal;
 Transferase; Tyrosine-protein kinase.
 FT SIGNAL 1 21 POTENTIAL.
 FT CHAIN 22 724 FIBROBLAST GROWTH FACTOR RECEPTOR.
 SQ SEQUENCE 724 AA; 80899 MW; 036D35842ED5EDB3 CRC64;

Query Match 90.7%; Score 97; DB 13; Length 724;
 Best Local Similarity 90.9%; Pred. No. 5.1e-08;
 Matches 20; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 HSQMAVHKLAKSIPLRRQVTVS 22
 DB 317 NSQLAVHKLAKSIPLRRQVTVS 338

RESULT 11
 Q90200 PRELIMINARY; PRT; 806 AA.
 AC Q90200;
 DT 01-DEC-2001 (TREMELrel. 19, Created)
 DT 01-DEC-2001 (TREMELrel. 19, Last sequence update)

DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
 DE Fibroblast growth factor receptor 1.
 GN FGFR1.
 OS Brachydanio rerio (Zebrafish) (Danio rerio).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
 OC Cyprinidae; Danio.
 OX NCBI_TaxID=7955;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Groth C., Scholpp S., Lardelli M., Brand M.;
 RT "Expression analysis of zebrafish fgfr1 isoforms";
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF389400; AAK64494.1; -
 DR ZFIN; ZDB-GENE-980526-255; fgfr1.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.
 DR GO; GO:0004872; F:receptor activity; IEA.
 DR GO; GO:0016740; F:transferase activity; IEA.
 DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
 DR InterPro; IPR007110; IG-like.
 DR InterPro; IPR003598; IG_c2.
 DR InterPro; IPR000719; Prot_kinase.
 DR InterPro; IPR001245; Tyr_pkinase.
 DR InterPro; IPR008266; Tyr_pkinase_AS.
 DR Pfam; PF00047; IG; 3.
 DR PRINTS; PS00069; pkinase; 1.
 DR SMART; SM00408; IGC2; 3.
 DR SMART; SM00219; TyrKc; 1.
 DR PROSITE; PS00835; IG LIKE; 3.
 DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
 DR PROSITE; PS00111; PROTEIN KINASE DOM; 1.
 DR PROSITE; PS00109; PROTEIN KINASE TYR; 1.
 KW ATP-binding; Immunoglobulin domain; Kinase; Receptor; Transferase;
 KW Tyrosine-protein kinase.
 SQ SEQUENCE 806 AA; 90489 MW; 056AF51FEAB0FBF9 CRC64;

Query Match 90.7%; Score 87; DB 13; Length 806;
 Best Local Similarity 90.9%; Pred. No. 5.7e-08;
 Matches 20; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 HSQAVHKLAKSIPLRRQVTVS 22
 DB 392 NSQLAVHKLAKSIPLRRQVTVS 413

RESULT 12

Q1897
 ID Q91897 PRELIMINARY; PRT; 814 AA.
 AC Q91897;
 DT 01-NOV-1996 (TREMELrel. 01, Created)
 DT 01-NOV-1996 (TREMELrel. 01, Last sequence update)
 DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
 DE Fibroblast growth factor receptor precursor.
 GN X1FGFR OR E59.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
 OC Xenopodinae; Xenopus.
 OX NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Priesel R.E., David I.B.;
 RL Submitted (AUG-1991) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE OF 622-677 FROM N.A.
 RX MEDLINE=95383727; PubMed=7655077;
 RA Brandl A.W., Kirschner M.W.;
 RT "Molecular cloning of tyrosine kinases in the early Xenopus embryo;
 RT identification of E2k-related genes expressed in cranial neural crest
 RT cells of the second (hyoid) arch."

RL Dev. Dyn. 203:119-140(1995).
 DR EMBL; M55163; AAA49990.1; -
 DR EMBL; U11723; AAA91286.1; -
 DR PIR; A39752; A39752.
 DR HSSP; P11362; IFGK.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.
 DR GO; GO:0016740; F:transferase activity; IEA.
 DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
 DR InterPro; IPR007110; IG-like.
 DR InterPro; IPR003598; IG_c2.
 DR InterPro; IPR000719; Prot_kinase.
 DR InterPro; IPR001245; Tyr_pkinase.
 DR InterPro; IPR008266; Tyr_pkinase_AS.
 DR Pfam; PF00047; IG; 3.
 DR Pfam; PF00069; pkinase; 1.
 DR PRINTS; PR00109; TYRKINASE.
 DR ProDom; PD000001; Prot_kinase; 1.
 DR SMART; SM00408; IGC2; 3.
 DR SMART; SM00219; TyrKc; 1.
 DR PROSITE; PS00835; IG LIKE; 3.
 DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
 DR PROSITE; PS00111; PROTEIN KINASE DOM; 1.
 DR PROSITE; PS00109; PROTEIN KINASE TYR; 1.
 KW ATP-binding; Immunoglobulin domain; Kinase; Signal; Transferase;
 KW Tyrosine-protein kinase.
 FT SIGNAL 1 21
 FT CHAIN 22 814
 SQ SEQUENCE 814 AA; 90681 MW; B47CE05F0A8F962A CRC64;

Query Match 90.7%; Score 97; DB 13; Length 814;
 Best Local Similarity 90.9%; Pred. No. 5.8e-08;
 Matches 20; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 HSQAVHKLAKSIPLRRQVTVS 22
 DB 405 NSQLAVHKLAKSIPLRRQVTVS 426

RESULT 13

Q63827
 ID Q63827 PRELIMINARY; PRT; 729 AA.
 AC Q63827;
 DT 01-NOV-1996 (TREMELrel. 01, Created)
 DT 01-NOV-1996 (TREMELrel. 01, Last sequence update)
 DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
 DE Fibroblast growth factor receptor 1 beta-isoform.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93158788; PubMed=8381605;
 RA Kim E.G., Kwon H.M., Burrow C.R., Ballermann B.J.;
 RT "Expression of rat fibroblast growth factor receptor 1 as three
 RT splicing variants during kidney development.";
 RL Am. J. Physiol. 264:F66-F73(1993).
 DR EMBL; S54008; AAB54274.1; -
 DR PIR; A56795; A56795.
 DR HSSP; P11362; IFGK.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.
 DR GO; GO:0016740; F:transferase activity; IEA.
 DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
 DR InterPro; IPR007110; IG-like.
 DR InterPro; IPR003598; IG_c2.
 DR InterPro; IPR000719; Prot_kinase.
 DR InterPro; IPR001245; Tyr_pkinase.
 DR InterPro; IPR008266; Tyr_pkinase_AS.
 DR Pfam; PF00047; IG; 2.
 DR Pfam; PF00069; pkinase; 1.
 DR PRINTS; PR00109; TYRKINASE.

```
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00408; IGC2; 2.
DR SMART; SM00219; TYRK; 1.
DR PROSITE; PS00835; IG_LIKE; 2.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE; PS00109; IMMUNOGLOBULIN_DOMAIN; Kinase; Transferase;
KW Tyrosine-protein kinase.
KW ATP-binding.
SQ SEQUENCE 729 AA; 81627 MW; C19DFDAFFD5BA0BF CRC64;

Query Match 88.8%; Score 95; DB 11; Length 729;
Best Local Similarity 95.0%; Pred. No. 1.1e-07;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 H5QMAVHKLAKSIPLRRQTV 20
   |||||
Db 318 H5QMAVHKLAKSIPLRRQVS 337

RESULT 14
ID Q80020 PRELIMINARY; PRT; 756 AA.
AC Q80020;
DT 01-JUN-2003 (TRENBLrel. 24, Created)
DT 01-JUN-2003 (TRENBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Fibroblast growth factor receptor 1 isoform 3.
FGFR1.
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RA Groth C., Scholpp S., Brand M., Lardelli M.;
RT "developmental expression of zebrafish fgfr1";
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY197499; AAC45659.1; -.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR003599; IG.
DR InterPro; IPR007110; IG-like.
DR InterPro; IPR003598; IG_c2.
DR InterPro; IPR00719; Prot_kinase.
DR InterPro; IPR002290; Ser_Ehr_pkinase.
DR InterPro; IPR001245; Tyr_pkinase.
DR InterPro; IPR008266; Tyr_pkinase_AS.
DR Pfam; PF00047; Ig; 3.
DR Pfam; PF00069; pkinase; 1.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00409; IG; 3.
DR SMART; SM00408; IGC2; 3.
DR SMART; SM00220; S_TKC; 1.
DR SMART; SM00219; TYRK; 1.
DR PROSITE; PS00835; IG_LIKE; 2.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
KW Receptor.
SQ SEQUENCE 756 AA; 85059 MW; 2B898370C7402507 CRC64;

Query Match 83.2%; Score 89; DB 13; Length 756;
Best Local Similarity 85.7%; Pred. No. 1.3e-06;
Matches 18; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 H5QMAVHKLAKSIPLRRQTV 21
   :|||
Db 392 NSQLAVHKLAKSIPLRRQVS 412

Search completed: August 2, 2004, 09:31:16
Job time : 10.8421 secs
```

Query Match 83.2%; Score 89; DB 13; Length 804; Best Local Similarity 85.7%; Pred. No. 1.4e-06; Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 H5QMAVHKLAKSIPLRRQTV 21

Db 392 NSQLAVHKLAKSIPLRRQVS 412

Search completed: August 2, 2004, 09:31:16
Job time : 10.8421 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 2, 2004, 09:27:44 ; Search time 4.19737 Seconds
(without alignments)
270.591 Million cell updates/sec

Title: US-09-757-415A-3

Perfect score: 107

Sequence: 1 H5QMAVHKLAKSIPLRRQVTVS 22

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA:*
1: /cgn2_6/prodata/2/iaa/5A COMB.pap.*
2: /cgn2_6/prodata/2/iaa/5B COMB.pap.*
3: /cgn2_6/prodata/2/iaa/6A COMB.pap.*
4: /cgn2_6/prodata/2/iaa/6B COMB.pap.*
5: /cgn2_6/prodata/2/iaa/PCUTS COMB.pap.*
6: /cgn2_6/prodata/2/iaa/backfiles1.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	107	100.0	729	1	US-07-640-029-3
2	107	100.0	731	1	US-07-921-807B-5
3	107	100.0	731	1	US-08-441-944A-5
4	107	100.0	733	1	US-08-439-992A-3
5	107	100.0	733	1	US-07-640-029-4
6	107	100.0	733	1	US-07-921-807B-6
7	107	100.0	733	1	US-08-441-944A-6
8	107	100.0	733	1	US-08-439-992A-4
9	107	100.0	816	1	US-07-640-029-1
10	107	100.0	817	1	US-07-640-029-2
11	107	100.0	820	1	US-07-921-807B-3
12	107	100.0	820	1	US-08-441-944A-3
13	107	100.0	820	1	US-08-166-717D-6
14	107	100.0	820	3	US-08-439-992A-1
15	107	100.0	822	1	US-07-997-133-1
16	107	100.0	822	1	US-07-921-807B-4
17	107	100.0	822	1	US-08-459-996-2
18	107	100.0	822	1	US-08-441-944A-4
19	107	100.0	822	2	US-08-451-822A-12
20	107	100.0	822	3	US-08-439-992A-2
21	107	100.0	822	4	US-08-323-430-12
22	83	77.6	643	1	US-08-471-570-6
23	83	77.6	769	1	US-08-471-570-8
24	83	77.6	821	2	US-08-451-822A-13
25	83	77.6	821	4	US-08-323-430-13
26	81	75.7	729	1	US-08-070-165F-6
27	81	75.7	729	2	US-08-885-418-6

Sequence 10, Appl
Sequence 10, Appl
Sequence 4, Appl
Sequence 10, Appl
Sequence 13345, A
Sequence 17924, A
Sequence 5, Appl
Sequence 7, Appl
Sequence 7, Appl
Sequence 3143, Ap
Sequence 30297, A
Sequence 32541, A
Sequence 8057, Ap
Sequence 16, Appl
Sequence 27, Appl
Sequence 26045, A
Sequence 23212, A
Sequence 478, App

ALIGNMENTS

RESULT 1

US-07-640-029-3
; Sequence 3, Application US/076400029
; Patent No. 5229501
; GENERAL INFORMATION:
; APPLICANT: Kiefer, Michael C.
; APPLICANT: Valenzuela, Pablo D.T.
; APPLICANT: Barr, Phillip J.
; TITLE OF INVENTION: Expression and Use of Human Fibroblast
; TITLE OF INVENTION: Growth Factor Receptor
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation
; STREET: 4560 Horton Street
; CITY: Emeryville
; STATE: California
; COUNTRY: USA
; ZIP: 94608
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/640,029
; FILING DATE: 19910111
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: McClung, Barbara G.
; REGISTRATION NUMBER: 33,113
; REFERENCE/DOCKET NUMBER: CH-165
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 510-601-2708
; TELEFAX: 510-655-3542
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 729 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-07-640-029-3

Query Match 100.0%; Score 107; DB 1; Length 729;
Best Local Similarity 100.0%; Pred. No. 9.1e-10;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 H5QMAVHKLAKSIPLRRQVTVS 22
Db 316 H5QMAVHKLAKSIPLRRQVTVS 337

RESULT 2
US-07-921-807B-5
; Sequence 5, Application US/07921807B
; Patent No. 5474914
; GENERAL INFORMATION:
; APPLICANT: SPAETE, RICHARD
; TITLE OF INVENTION: METHOD OF INCREASING EXPRESSION
; TITLE OF INVENTION: OF VIRAL PROTEINS
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CHIRON CORPORATION
; STREET: 4560 Horton Street - R440
; CITY: Emeryville
; STATE: CA
; COUNTRY: USA
; ZIP: 94608-2916
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/921.807B
; FILING DATE: 29-SEP-1992
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: MCCLUNG, BARBARA G.
; REGISTRATION NUMBER: 33,113
; REFERENCE/DOCKET NUMBER: 0209.001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (510) 601-2708
; TELEFAX: (510) 655-3542
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 731 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-921-807B-5

Query Match 100.0%; Score 107; DB 1; Length 731;
Best Local Similarity 100.0%; Pred. No. 9.2e-10;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HSQMAVHKLAKSIPLRRQVTVS 22
Db 318 HSQMAVHKLAKSIPLRRQVTVS 339

RESULT 3
US-08-441-944A-5
; Sequence 5, Application US/08441944A
; Patent No. 5767250
; GENERAL INFORMATION:
; APPLICANT: SPAETE, RICHARD
; TITLE OF INVENTION: METHOD OF INCREASING EXPRESSION
; TITLE OF INVENTION: OF VIRAL PROTEINS
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CHIRON CORPORATION
; STREET: 4560 Horton Street - R440
; CITY: Emeryville
; STATE: CA
; COUNTRY: USA
; ZIP: 94608-2916
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/441,944A
; FILING DATE: 16-MAY-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/921,807
; FILING DATE: 29-SEP-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: MCCLUNG, BARBARA G.
; REGISTRATION NUMBER: 33,113
; REFERENCE/DOCKET NUMBER: 0209.001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (510) 601-2708
; TELEFAX: (510) 655-3542
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 731 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-441-944A-5

Query Match 100.0%; Score 107; DB 1; Length 731;
Best Local Similarity 100.0%; Pred. No. 9.2e-10;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HSQMAVHKLAKSIPLRRQVTVS 22
Db 318 HSQMAVHKLAKSIPLRRQVTVS 339

RESULT 4
US-08-439-992A-3
; Sequence 3, Application US/08439992A
; Patent No. 6255454
; GENERAL INFORMATION:
; APPLICANT: Kiefer, Michael C.
; APPLICANT: Pablo, Valenzuela D.T.
; APPLICANT: Philip, Barr J.
; TITLE OF INVENTION: Expression and Use of Human Fibroblast
; TITLE OF INVENTION: Receptor
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation
; STREET: 4560 Horton Street
; CITY: Emeryville
; STATE: CA
; COUNTRY: USA
; ZIP: 94608
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/439,992A
; FILING DATE: 12-MAY-1995
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Chung, Ling-Fong
; REGISTRATION NUMBER: 36,482
; REFERENCE/DOCKET NUMBER: 0165.004
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 510-923-2704
; TELEFAX: (510) 655-3542
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 731 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein

US-08-439-992A-3

Query Match 100.0%; Score 107; DB 3; Length 731;
Best Local Similarity 100.0%; Pred. No. 9.2e-10;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HSQMAVHKLAKSIPLRRQVTVS 22
Db 318 HSQMAVHKLAKSIPLRRQVTVS 339

RESULT 5
US-07-640-029-4
Sequence 4, Application US/07640029
Patent No. 5229501
GENERAL INFORMATION:
APPLICANT: Kiefer, Michael C.
APPLICANT: Valenzuela, Pablo D.T.
APPLICANT: Barr, Philip J.
TITLE OF INVENTION: Expression and Use of Human Fibroblast
TITLE OF INVENTION: Growth Factor Receptor
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Chiron Corporation
STREET: 4560 Horton Street
CITY: Emeryville
STATE: California
COUNTRY: USA
ZIP: 94608
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION NUMBER: US/07/640,029
FILING DATE: 19910111
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: McClung, Barbara G.
REGISTRATION NUMBER: 33,113
REFERENCE/DOCKET NUMBER: CH-165
TELECOMMUNICATION INFORMATION:
TELEPHONE: 510-601-2708
TELEFAX: 510-655-3542
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 733 amino acids
TYPE: AMINO ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-07-640-029-4

Query Match 100.0%; Score 107; DB 1; Length 733;
Best Local Similarity 100.0%; Pred. No. 9.2e-10;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HSQMAVHKLAKSIPLRRQVTVS 22
Db 320 HSQMAVHKLAKSIPLRRQVTVS 341

RESULT 6
US-07-921-807B-6
Sequence 6, Application US/07921807B
Patent No. 547914
GENERAL INFORMATION:
APPLICANT: SPAETE, RICHARD
TITLE OF INVENTION: METHOD OF INCREASING EXPRESSION
TITLE OF INVENTION: OF VIRAL PROTEINS
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:

ADDRESSEE: CHIRON CORPORATION
STREET: 4560 Horton Street - R440
CITY: Emeryville
STATE: CA
COUNTRY: USA
ZIP: 94608-2916
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA: US/07/921,807B
APPLICATION NUMBER: US/07/921,807B
FILING DATE: 29-SEP-1992
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: MCCLUNG, BARBARA G.
REGISTRATION NUMBER: 33,113
REFERENCE/DOCKET NUMBER: 0209,001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (510) 601-2708
TELEFAX: (510) 655-3542
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 733 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-921-807B-6

Query Match 100.0%; Score 107; DB 1; Length 733;
Best Local Similarity 100.0%; Pred. No. 9.2e-10;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HSQMAVHKLAKSIPLRRQVTVS 22
Db 320 HSQMAVHKLAKSIPLRRQVTVS 341

RESULT 7
US-08-441-944A-6
Sequence 6, Application US/08441944A
Patent No. 5767250
GENERAL INFORMATION:
APPLICANT: SPAETE, RICHARD
TITLE OF INVENTION: METHOD OF INCREASING EXPRESSION
TITLE OF INVENTION: OF VIRAL PROTEINS
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: CHIRON CORPORATION
STREET: 4560 Horton Street - R440
CITY: Emeryville
STATE: CA
COUNTRY: USA
ZIP: 94608-2916
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/441,944A
FILING DATE: 16-MAY-1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/921,807
FILING DATE: 29-SEP-1992
ATTORNEY/AGENT INFORMATION:
NAME: MCCLUNG, BARBARA G.
REGISTRATION NUMBER: 33,113
REFERENCE/DOCKET NUMBER: 0209,001
TELECOMMUNICATION INFORMATION:

```

; TELEPHONE: (510) 601-2708
; TELEFAX: (510) 655-3542
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 733 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-441-944A-6

Query Match 100.0%; Score 107; DB 1; Length 733;
Best Local Similarity 100.0%; Pred. No. 9.2e-10;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HSQMAVHKLAKSIPLRRQVTVS 22
DB 320 HSQMAVHKLAKSIPLRRQVTVS 341

RESULT 8
US-08-439-992A-4
; Sequence 4, Application US/08439992A
; Patent No. 6255454
; GENERAL INFORMATION:
; APPLICANT: Kiefer, Michael C.
; APPLICANT: Pablo, Valenzuela D.T.
; TITLE OF INVENTION: Expression and Use of Human Fibroblast
; TITLE OF INVENTION: Receptor
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation
; STREET: 4560 Horton Street
; CITY: Emeryville
; STATE: CA
; COUNTRY: USA
; ZIP: 94608
; COMPUTER READABLE FORM:
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/439,992A
; FILING DATE: 12-MAY-1995
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Chung, Ling-Fong
; REGISTRATION NUMBER: 36,482
; REFERENCE/DOCKET NUMBER: 0165.004
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 510-923-2704
; TELEFAX: (510) 655-3542
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 733 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-439-992A-4

Query Match 100.0%; Score 107; DB 3; Length 733;
Best Local Similarity 100.0%; Pred. No. 9.2e-10;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HSQMAVHKLAKSIPLRRQVTVS 22
DB 320 HSQMAVHKLAKSIPLRRQVTVS 341

RESULT 9
```

```

US-07-640-029-1
; Sequence 1, Application US/07640029
; Patent No. 5229501
; GENERAL INFORMATION:
; APPLICANT: Kiefer, Michael C.
; APPLICANT: Valenzuela, Pablo D.T.
; APPLICANT: Barr, Philip J.
; TITLE OF INVENTION: Expression and Use of Human Fibroblast
; TITLE OF INVENTION: Growth Factor Receptor
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation
; STREET: 4560 Horton Street
; CITY: Emeryville
; STATE: California
; COUNTRY: USA
; ZIP: 94608
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/640,029
; FILING DATE: 19910111
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: McClung, Barbara G.
; REGISTRATION NUMBER: 33,113
; REFERENCE/DOCKET NUMBER: CH-165
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 510-601-2708
; TELEFAX: 510-655-3542
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 816 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-07-640-029-1

Query Match 100.0%; Score 107; DB 1; Length 816;
Best Local Similarity 100.0%; Pred. No. 1e-09;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HSQMAVHKLAKSIPLRRQVTVS 22
DB 404 HSQMAVHKLAKSIPLRRQVTVS 425

RESULT 10
US-07-640-029-2
; Sequence 2, Application US/07640029
; Patent No. 5229501
; GENERAL INFORMATION:
; APPLICANT: Kiefer, Michael C.
; APPLICANT: Valenzuela, Pablo D.T.
; APPLICANT: Barr, Philip J.
; TITLE OF INVENTION: Expression and Use of Human Fibroblast
; TITLE OF INVENTION: Growth Factor Receptor
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation
; STREET: 4560 Horton Street
; CITY: Emeryville
; STATE: California
; COUNTRY: USA
; ZIP: 94608
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
```

;; SOFTWARE: PatentIn Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/07/640,029
;; FILING DATE: 19910111
;; CLASSIFICATION: 530
;; ATTORNEY/AGENT INFORMATION:
;; NAME: McClung, Barbara G.
;; REGISTRATION NUMBER: 33,113
;; REFERENCE/DOCKET NUMBER: CH-165
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 510-601-2708
;; TELEFAX: 510-655-3542
;; INFORMATION FOR SEQ ID NO: 2:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 817 amino acids
;; TYPE: AMINO ACID
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
US-07-640-029-2

Query Match 100.0%; Score 107; DB 1; Length 817;
Best Local Similarity 100.0%; Pred. No. 1e-09;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HSQMAVHKLAKSIPLRRQVTVS 22
Db 409 HSQMAVHKLAKSIPLRRQVTVS 430

RESULT 11
US-07-921-807B-3
;; Sequence 3, Application US/07921807B
;; Patent No. 5474914
;; GENERAL INFORMATION:
;; APPLICANT: SPAETE, RICHARD
;; TITLE OF INVENTION: METHOD OF INCREASING EXPRESSION
;; TITLE OF INVENTION: OF VIRAL PROTEINS
;; NUMBER OF SEQUENCES: 20
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: CHIRON CORPORATION
;; STREET: 4560 Horton Street - R440
;; CITY: Emeryville
;; STATE: CA
;; COUNTRY: USA
;; ZIP: 94608-2916
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/07/921,807B
;; FILING DATE: 29-SEP-1992
;; CLASSIFICATION: 435
;; ATTORNEY/AGENT INFORMATION:
;; NAME: McCLUNG, BARBARA G.
;; REGISTRATION NUMBER: 33,113
;; REFERENCE/DOCKET NUMBER: 0209.001
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (510) 601-2708
;; TELEFAX: (510) 655-3542
;; INFORMATION FOR SEQ ID NO: 3:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 820 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
US-07-921-807B-3

Query Match 100.0%; Score 107; DB 1; Length 820;
Best Local Similarity 100.0%; Pred. No. 1e-09;

Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 HSQMAVHKLAKSIPLRRQVTVS 22
Db 407 HSQMAVHKLAKSIPLRRQVTVS 428

RESULT 12
US-08-441-944A-3
;; Sequence 3, Application US/08441944A
;; Patent No. 5767250
;; GENERAL INFORMATION:
;; APPLICANT: SPAETE, RICHARD
;; TITLE OF INVENTION: METHOD OF INCREASING EXPRESSION
;; TITLE OF INVENTION: OF VIRAL PROTEINS
;; NUMBER OF SEQUENCES: 20
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: CHIRON CORPORATION
;; STREET: 4560 Horton Street - R440
;; CITY: Emeryville
;; STATE: CA
;; COUNTRY: USA
;; ZIP: 94608-2916
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/441,944A
;; FILING DATE: 16-MAY-1995
;; CLASSIFICATION: 530
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 07/921,807
;; FILING DATE: 29-SEP-1992
;; ATTORNEY/AGENT INFORMATION:
;; NAME: McCLUNG, BARBARA G.
;; REGISTRATION NUMBER: 33,113
;; REFERENCE/DOCKET NUMBER: 0209.001
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (510) 601-2708
;; TELEFAX: (510) 655-3542
;; INFORMATION FOR SEQ ID NO: 3:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 820 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
US-08-441-944A-3

Query Match 100.0%; Score 107; DB 1; Length 820;
Best Local Similarity 100.0%; Pred. No. 1e-09;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HSQMAVHKLAKSIPLRRQVTVS 22
Db 407 HSQMAVHKLAKSIPLRRQVTVS 428

RESULT 13
US-08-166-717D-6
;; Sequence 6, Application US/08166717D
;; Patent No. 5789182
;; GENERAL INFORMATION:
;; APPLICANT: Yavon, Avner
;; APPLICANT: Ornitz, David M.
;; APPLICANT: Klagsbrun, Michael
;; APPLICANT: Leder, Philip
;; TITLE OF INVENTION: SYSTEM FOR ASSAYING BINDING
;; TITLE OF INVENTION: TO A HEPARIN-BINDING GROWTH
;; TITLE OF INVENTION: FACTOR RECEPTOR
;; NUMBER OF SEQUENCES: 6

```
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Clark & Ebling LLP
/ STREET: 176 Federal Street
/ CITY: Boston
/ STATE: Massachusetts
/ COUNTRY: U.S.A.
/ ZIP: 02110
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
/ COMPUTER: IBM COMPATIBLE
/ OPERATING SYSTEM: WINDOWS 95
/ SOFTWARE: WordPerfect (Version 7.0)
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/166.717D
/ FILING DATE: 12/14/93
/ CLASSIFICATION: 435
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 07/631,717
/ FILING DATE: 12/20/90
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Kristina Bieker-Brady
/ REGISTRATION NUMBER: 39,109
/ REFERENCE/DOCKET NUMBER: 00383/017002
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (617) 723-4123
/ TELEFAX: (617) 723-8962
/ TELEX:
/ INFORMATION FOR SEQ ID NO: 6:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 820
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ US-08-166-717D-6
/
/ Query Match 100.0%; Score 107; DB 1; Length 820;
/ Best Local Similarity 100.0%; Pred. No. 1e-09;
/ Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
/
/ QY 1 HSQMAVHKLAKSIPLRRQVTWS 22
/ Db 407 HSQMAVHKLAKSIPLRRQVTWS 428
/
/ RESULT 14
/ US-08-439-992A-1
/ Sequence 1, Application US/08439992A
/ Patent No. 6255454
/ GENERAL INFORMATION:
/ APPLICANT: Kiefer, Michael C.
/ APPLICANT: Pablo, Valenzuela D.T.
/ APPLICANT: Philip, Barr J.
/ TITLE OF INVENTION: Expression and Use of Human Fibroblast
/ NUMBER OF SEQUENCES: 12
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Chiron Corporation
/ STREET: 4560 Horton Street
/ CITY: Emeryville
/ STATE: CA
/ COUNTRY: USA
/ ZIP: 94608
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/439,992A
/ FILING DATE: 12-MAY-1995
/ CLASSIFICATION: 530
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Chung, Ling-Fong
/
/ Query Match 100.0%; Score 107; DB 1; Length 822;
/ Best Local Similarity 100.0%; Pred. No. 1e-09;
/
/ /
/ REGISTRATION NUMBER: 36,482
/ REFERENCE/DOCKET NUMBER: 0165.004
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 510-923-2704
/ TELEFAX: (510) 655-3542
/ INFORMATION FOR SEQ ID NO: 1:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 820 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ US-08-439-992A-1
/
/ Query Match 100.0%; Score 107; DB 3; Length 820;
/ Best Local Similarity 100.0%; Pred. No. 1e-09;
/ Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
/
/ QY 1 HSQMAVHKLAKSIPLRRQVTWS 22
/ Db 407 HSQMAVHKLAKSIPLRRQVTWS 428
/
/ RESULT 15
/ US-07-997-133-1
/ Sequence 1, Application US/07997133
/ Patent No. 5288855
/ GENERAL INFORMATION:
/ APPLICANT: Bergonzoni, Laura
/ APPLICANT: Mazue, Guy
/ APPLICANT: Isacchi, Antonella
/ APPLICANT: Roncucci, Romeo
/ APPLICANT: Sarmientos, Paolo
/ TITLE OF INVENTION: Extracellular Form of the Human
/ NUMBER OF SEQUENCES: 8
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
/ STREET: 1735 Jefferson Davis Highway, Fourth Floor
/ CITY: Arlington
/ STATE: Virginia
/ ZIP: 22202
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/07/997,133
/ FILING DATE: 28-DEC-1992
/ CLASSIFICATION: 530
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US/07/642,755
/ FILING DATE: 18-JAN-1991
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Oblon, No. 5288855man F.
/ REGISTRATION NUMBER: 24,618
/ REFERENCE/DOCKET NUMBER: 769-226-0
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (703) 521-4500
/ TELEFAX: (703) 486-2347
/ INFORMATION FOR SEQ ID NO: 1:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 822 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ US-07-997-133-1
/
/ Query Match 100.0%; Score 107; DB 1; Length 822;
/ Best Local Similarity 100.0%; Pred. No. 1e-09;
```

Mon Aug 9 09:19:07 2004

	Matches	22;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
Qy	1	HSCMAVHKLAKSIPLRRQVTVS	22							
Db	409	HSCMAVHKLAKSIPLRRQVTVS	430							

Search completed: August 2, 2004, 09:32:29
Job time : 5.19737 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 2, 2004, 09:31:19 ; Search time 11.7237 Seconds
(without alignments)
588.640 Million cell updates/sec

Title: US-09-757-415A-3

Perfect score: 107

Sequence: 1 HSQMAVHKLAKSIPLRRQVTVS 22

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1291235 seqs, 313682936 residues

Total number of hits satisfying chosen parameters: 1291235

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA:

1: /cgn2_6/prodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/prodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/prodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/prodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/prodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/prodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/prodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/prodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/prodata/2/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/prodata/2/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/prodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/prodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/prodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/prodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/prodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/prodata/2/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/prodata/2/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/prodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	DB ID	Description
1	107	100.0	22	US-09-757-415A-3
2	107	100.0	702	US-09-805-020-47
3	107	100.0	735	US-10-307-817-6
4	107	100.0	764	US-09-925-302-714
5	107	100.0	764	US-09-925-302-714
6	107	100.0	820	US-10-302-812-42
7	107	100.0	822	US-09-757-415A-2
8	107	100.0	822	US-10-307-817-8
9	107	100.0	822	US-10-304-041-2
10	107	100.0	822	US-10-394-322A-26
11	107	100.0	824	US-10-307-817-4
12	101	94.4	451	US-10-087-192-381
13	101	94.4	609	US-10-087-192-384
14	83	77.6	821	US-10-394-322A-27
15	83	77.6	821	US-10-302-812-44

83 77.6 822 16 US-10-648-593-254 Sequence 254, App
77 72.0 471 12 US-10-087-192-1134 Sequence 1134, App
64 59.8 20 14 US-10-199-820-247 Sequence 247, App
47 43.9 4370 16 US-10-408-765A-1267 Sequence 1267, App
47 43.9 4393 12 US-10-231-956A-366 Sequence 366, App
21 44 41.1 63 16 US-10-437-963-145357 Sequence 145357,
22 44 41.1 95 12 US-10-424-599-227789 Sequence 227789,
23 44 41.1 194 9 US-09-738-626-3731 Sequence 3731, App
24 44 41.1 214 12 US-10-424-599-227795 Sequence 227795,
25 44 41.1 234 12 US-10-424-599-227788 Sequence 227788,
26 44 41.1 1520 14 US-10-017-161-718 Sequence 718, App
27 44 41.1 1520 15 US-10-292-798-630 Sequence 630, App
28 43.5 40.7 60 16 US-10-437-963-115591 Sequence 115591,
29 43 40.2 40 12 US-10-424-599-203535 Sequence 203535,
30 43 40.2 152 12 US-10-424-599-150884 Sequence 150884,
31 43 40.2 152 16 US-10-437-963-116361 Sequence 116361,
32 43 40.2 207 16 US-10-437-963-155843 Sequence 155843,
33 43 40.2 277 12 US-10-282-122A-51737 Sequence 51737, A
34 43 40.2 277 12 US-10-358-448A-8 Sequence 8, Appli
35 43 40.2 339 15 US-10-369-493-16476 Sequence 16476, A
36 43 40.2 446 12 US-10-292-122A-77108 Sequence 77108, A
37 43 40.2 778 16 US-10-437-963-118289 Sequence 118289,
38 42.5 39.7 187 12 US-10-425-114-43393 Sequence 43393, A
39 42 39.3 202 12 US-09-800-095A-70 Sequence 70, Appl
40 42 39.3 342 15 US-10-108-260A-4204 Sequence 4204, A
41 42 39.3 374 15 US-10-369-493-19707 Sequence 19707, A
42 42 39.3 377 12 US-10-282-122A-63371 Sequence 63371, A
43 42 39.3 382 12 US-10-425-114-38931 Sequence 38931, A
44 42 39.3 418 16 US-10-461-194-118 Sequence 118, App

ALIGNMENTS

RESULT 1

US-09-757-415A-3
; Sequence 3, Application US/09757415A
; Publication No. US20030040612A1
; GENERAL INFORMATION:
; APPLICANT: Zhou, Ming-Ming
; TITLE OF INVENTION: Methods of Identifying Modulators of the FGF Receptor
; FILE REFERENCE: 2459-1-002N
; CURRENT APPLICATION NUMBER: US/09/757,415A
; CURRENT FILING DATE: 2001-01-09
; PRIOR APPLICATION NUMBER: 60/175867
; PRIOR FILING DATE: 2000-01-12
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 22
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-757-415A-3

Query Match 100.0%; Score 107; DB 10; Length 22;
Best Local Similarity 100.0%; Pred. No. 4.9e-11;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HSQMAVHKLAKSIPLRRQVTVS 22

DB 1 HSQMAVHKLAKSIPLRRQVTVS 22

RESULT 2

US-09-805-020-47
; Sequence 47, Application US/09805020
; Publication No. US20020086384A1
; GENERAL INFORMATION:
; APPLICANT: LEVINE, Zurit
; TITLE OF INVENTION: SPLICE VARIANTS OF ONCOGENES
; FILE REFERENCE: 2786-0168P
; CURRENT APPLICATION NUMBER: US/09/805,020

; CURRENT FILING DATE: 2001-03-13
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 47
; LENGTH: 702
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-805-020-47

Query Match 100.0%; Score 107; DB 12; Length 702;
Best Local Similarity 100.0%; Pred. No. 2.5e-09;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HSQMAVHKLAISIPLRRQVTVS 22
|||||
Db 409 HSQMAVHKLAISIPLRRQVTVS 430

RESULT 3

US-10-307-817-6
; Sequence 6, Application US/10307817
; Publication No. US20040058338A1
; GENERAL INFORMATION:
; APPLICANT: Agere et al.
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 21402-502C
; CURRENT APPLICATION NUMBER: US/10/307,817
; CURRENT FILING DATE: 2002-12-02
; NUMBER OF SEQ ID NOS: 682
; SOFTWARE: CuraseqList version 0.1
; SEQ ID NO 6
; LENGTH: 735
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-307-817-6

Query Match 100.0%; Score 107; DB 12; Length 735;
Best Local Similarity 100.0%; Pred. No. 2.7e-09;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HSQMAVHKLAISIPLRRQVTVS 22
|||||
Db 322 HSQMAVHKLAISIPLRRQVTVS 343

RESULT 4

US-09-925-302-714
; Sequence 714, Application US/09925302
; Patent No. US20020044941A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA104
; CURRENT APPLICATION NUMBER: US/09/925,302
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05918
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 896
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 714
; LENGTH: 764
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (44)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (62)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (81)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (125)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (725)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-302-714

Query Match 100.0%; Score 107; DB 12; Length 764;
Best Local Similarity 100.0%; Pred. No. 2.8e-09;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HSQMAVHKLAISIPLRRQVTVS 22
|||||
Db 351 HSQMAVHKLAISIPLRRQVTVS 372

RESULT 6

US-10-302-812-42
; Sequence 42, Application US/10302812
; Publication No. US20040087016A1
; GENERAL INFORMATION:
; APPLICANT: Keating et al.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR CELL DIFFERENTIATION AND

; LOCATION: (81)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (125)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (725)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-302-714

Query Match 100.0%; Score 107; DB 9; Length 764;
Best Local Similarity 100.0%; Pred. No. 2.8e-09;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HSQMAVHKLAISIPLRRQVTVS 22
|||||
Db 351 HSQMAVHKLAISIPLRRQVTVS 372

RESULT 5

US-09-925-302-714
; Sequence 714, Application US/09925302
; Publication No. US20030064072A9
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA104
; CURRENT APPLICATION NUMBER: US/09/925,302
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05918
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 896
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 714
; LENGTH: 764
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (44)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (62)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (81)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (125)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (725)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-302-714

Query Match 100.0%; Score 107; DB 12; Length 764;
Best Local Similarity 100.0%; Pred. No. 2.8e-09;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HSQMAVHKLAISIPLRRQVTVS 22
|||||
Db 351 HSQMAVHKLAISIPLRRQVTVS 372

RESULT 6

US-10-302-812-42
; Sequence 42, Application US/10302812
; Publication No. US20040087016A1
; GENERAL INFORMATION:
; APPLICANT: Keating et al.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR CELL DIFFERENTIATION AND

; TITLE OF INVENTION: TISSUE REGENERATION
; FILE REFERENCE: HYDR-P02-004
; CURRENT APPLICATION NUMBER: US/10/302,812
; CURRENT FILING DATE: 2002-11-21
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 42
; LENGTH: 820
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-302-812-42

Query Match 100.0%; Score 107; DB 16; Length 820;
Best Local Similarity 100.0%; Pred. No. 3e-09;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HSOQAVHKLAKSIPLRRQVTVS 22
Db 407 HSOQAVHKLAKSIPLRRQVTVS 428

RESULT 7
US-09-757-415A-2
; Sequence 2, Application US/09757415A
; Publication No. US20030040612A1
; GENERAL INFORMATION:
; APPLICANT: Zhou, Ming-Ming
; TITLE OF INVENTION: Methods of Identifying Modulators of the FGF Receptor
; FILE REFERENCE: 2459-1-002N
; CURRENT APPLICATION NUMBER: US/09/757,415A
; CURRENT FILING DATE: 2001-01-09
; PRIOR APPLICATION NUMBER: 60/175867
; PRIOR FILING DATE: 2000-01-12
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 822
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-757-415A-2

Query Match 100.0%; Score 107; DB 10; Length 822;
Best Local Similarity 100.0%; Pred. No. 3e-09;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HSOQAVHKLAKSIPLRRQVTVS 22
Db 409 HSOQAVHKLAKSIPLRRQVTVS 430

RESULT 8
US-10-307-817-8
; Sequence 8, Application US/10307817
; Publication No. US20040058338A1
; GENERAL INFORMATION:
; APPLICANT: Agee et al.
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 21402-502C
; CURRENT APPLICATION NUMBER: US/10/307,817
; CURRENT FILING DATE: 2002-12-02
; NUMBER OF SEQ ID NOS: 682
; SOFTWARE: Curaseq1 version 0.1
; SEQ ID NO 8
; LENGTH: 822
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-307-817-8

Query Match 100.0%; Score 107; DB 12; Length 822;
Best Local Similarity 100.0%; Pred. No. 3e-09;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HSOQAVHKLAKSIPLRRQVTVS 22

Db 409 HSOQAVHKLAKSIPLRRQVTVS 430

RESULT 9
US-10-204-041-2
; Sequence 2, Application US/10204041
; Publication No. US20030176443A1
; GENERAL INFORMATION:
; APPLICANT: STEIN-GERLACH, MATTHIAS
; APPLICANT: SALASSIDIS, KONSTANTINOS
; APPLICANT: BACHER, GERALD
; APPLICANT: MULLER, STEFAN
; TITLE OF INVENTION: Pyridylpyrimidine Derivatives as Effective Compounds Against Prior
; FILE REFERENCE: AXM-007.1P US
; CURRENT APPLICATION NUMBER: US/10/204,041
; CURRENT FILING DATE: 2002-08-16
; PRIOR APPLICATION NUMBER: EP 0111858.5
; PRIOR FILING DATE: 2001-05-16
; PRIOR APPLICATION NUMBER: PCT/EP02/05420
; PRIOR FILING DATE: 2002-05-16
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 822
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-204-041-2

Query Match 100.0%; Score 107; DB 14; Length 822;
Best Local Similarity 100.0%; Pred. No. 3e-09;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HSOQAVHKLAKSIPLRRQVTVS 22
Db 409 HSOQAVHKLAKSIPLRRQVTVS 430

RESULT 10
US-10-394-322A-26
; Sequence 26, Application US/10394322A
; Publication No. US2003023291A1
; GENERAL INFORMATION:
; APPLICANT: SUNESIS PHARMACEUTICALS, INC.
; APPLICANT: Prescott, John C.
; TITLE OF INVENTION: IDENTIFICATION OF KINASE INHIBITORS
; FILE REFERENCE: 39750-0006 US
; CURRENT APPLICATION NUMBER: US/10/394,322A
; CURRENT FILING DATE: 2003-03-20
; PRIOR APPLICATION NUMBER: US 60/366,892
; PRIOR FILING DATE: 2002-03-21
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 26
; LENGTH: 822
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-394-322A-26

Query Match 100.0%; Score 107; DB 15; Length 822;
Best Local Similarity 100.0%; Pred. No. 3e-09;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HSOQAVHKLAKSIPLRRQVTVS 22
Db 409 HSOQAVHKLAKSIPLRRQVTVS 430

RESULT 11
US-10-307-817-4
; Sequence 4, Application US/10307817
; Publication No. US20040058338A1

```
; GENERAL INFORMATION:
; APPLICANT: Agee et al.
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 21402-502C
; CURRENT APPLICATION NUMBER: US/10/307,817
; CURRENT FILING DATE: 2002-12-02
; NUMBER OF SEQ ID NOS: 682
; SOFTWARE: CuraseqList version 0.1
; SEQ ID NO 4
; LENGTH: 824
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-307-817-4

Query Match      100.0%; Score 107; DB 12; Length 824;
Best Local Similarity 100.0%; Pred. No. 3e-09;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HSQMAVHKLAKSIPLRRQVTVS 22
   |||||
Db 411 HSQMAVHKLAKSIPLRRQVTVS 432

RESULT 12
US-10-087-192-381
; Sequence 381, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: CANCER
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR FILING DATE: 2000-12-22
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 381
; LENGTH: 451
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-087-192-381

Query Match      94.4%; Score 101; DB 12; Length 451;
Best Local Similarity 95.5%; Pred. No. 1.6e-08;
Matches 21; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 HSQMAVHKLAKSIPLRRQVTVS 22
   |||||
Db 428 HSQMAVHKLAKSIPLRRQVTVS 449

RESULT 13
US-10-087-192-384
; Sequence 384, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: CANCER
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR FILING DATE: 2000-12-22
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
```

```
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 384
; LENGTH: 609
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-087-192-384

Query Match      94.4%; Score 101; DB 12; Length 609;
Best Local Similarity 95.5%; Pred. No. 2.2e-08;
Matches 21; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 HSQMAVHKLAKSIPLRRQVTVS 22
   |||||
Db 586 HSQMAVHKLAKSIPLRRQVTVS 607

RESULT 14
US-10-394-322A-27
; Sequence 27, Application US/10394322A
; Publication No. US200302391A1
; GENERAL INFORMATION:
; APPLICANT: SUNESIS PHARMACEUTICALS, INC.
; APPLICANT: Prescott, John C.
; TITLE OF INVENTION: IDENTIFICATION OF KINASE INHIBITORS
; FILE REFERENCE: 39750-0006 US
; CURRENT APPLICATION NUMBER: US/10/394,322A
; CURRENT FILING DATE: 2003-03-20
; PRIOR APPLICATION NUMBER: US 60/366,892
; PRIOR FILING DATE: 2002-03-21
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 27
; LENGTH: 821
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-394-322A-27

Query Match      77.6%; Score 83; DB 15; Length 821;
Best Local Similarity 85.7%; Pred. No. 3.7e-05;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 SQMAVHKLAKSIPLRRQVTVS 22
   |||||
Db 411 SQMAVHKLAKSIPLRRQVTVS 431

RESULT 15
US-10-302-812-44
; Sequence 44, Application US/10302812
; Publication No. US20040087016A1
; GENERAL INFORMATION:
; APPLICANT: Keating et al.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR CELL DIFFERENTIATION AND
; TITLE OF INVENTION: TISSUE REGENERATION
; FILE REFERENCE: HYDR-P02-004
; CURRENT APPLICATION NUMBER: US/10/302,812
; CURRENT FILING DATE: 2002-11-21
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 44
; LENGTH: 821
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-302-812-44

Query Match      77.6%; Score 83; DB 16; Length 821;
Best Local Similarity 85.7%; Pred. No. 3.7e-05;
Matches 18; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 SQMAVHKLAKSIPLRRQVTVS 22
   |||||
Db 411 SQMAVHKLAKSIPLRRQVTVS 431
```

Search completed: August 2, 2004, 09:40:46
Job time : 11.7237 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model
Run on: August 8, 2004, 19:45:10 ; Search time 806.618 Seconds
(without alignments)
1182.153 Million cell updates/sec

Title: US-09-757-415A-3
Perfect score: 107
Sequence: 1 H5QMAVHKLAKSIPLRRQVTVS 22

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgr2 1/USPTO.spool/US09757415/runat_02082004_101230_6685/app_query.fasta_1.526
-DB=GenEmbl -QFWT=fastap -SUPTX=p2n.rge -MINMATCH=0.1 -LOOCL=0 -LOOEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blsum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=ptt -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09757415 @CGN 1.1 3851 @runat_02082004_101230_6685 -NCPU=3
-NO MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSFLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPEXT=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl.*
1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.em.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.sy.*
13: gb.un.*
14: gb.vi.*
15: em.ba.*
16: em.fun.*
17: em.hum.*
18: em.in.*
19: em.mu.*
20: em.om.*
21: em.or.*
22: em.ov.*
23: em.pat.*
24: em.ph.*
25: em.pl.*
26: em.ro.*
27: em.sts.*
28: em.un.*

29: em.vi.*
30: em.htg_hum.*
31: em.htg_inv.*
32: em.htg_other.*
33: em.htg_mus.*
34: em.htg_pln.*
35: em.htg_rdt.*
36: em.htg_mam.*
37: em.htg_vrt.*
38: em_sy.*
39: em_htgo_hum.*
40: em_htgo_mus.*
41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	107	100.0	1608	9	HUMHBGFC M63889 Human hepar
2	107	100.0	1974	9	HSFLGMR Y09665 Human flg (
3	107	100.0	2259	10	MMU23445 U23445 Mus musculu
4	107	100.0	2360	6	AR380743 AR380743 Sequence
5	107	100.0	2360	9	HSFGR2IG X57122 Human mRNA
6	107	100.0	2366	9	HSFGR2IGA X57119 Human mRNA
7	107	100.0	2469	6	A29216 A29216 H.sapiens D
8	107	100.0	2469	6	I66313 I66313 Sequence 1
9	107	100.0	2469	6	AX481457 AX481457 Sequence
10	107	100.0	2469	10	RATFGFRL D12498 Rat mRNA fo
11	107	100.0	2499	10	MUSFGFBC M65053 Mouse fibro
12	107	100.0	2526	6	BD016930 BD016930 Antisense
13	107	100.0	2526	10	MMU2324 U23234 Mus musculu
14	107	100.0	2526	10	MUSBFGFR M28988 Mouse basic
15	107	100.0	2625	10	MMF6F X51893 Mouse fms-1
16	107	100.0	2627	9	HSFGR3IG X57120 Human mRNA
17	107	100.0	2633	9	HSFGR3IGA X57121 Human mRNA
18	107	100.0	2662	6	AX601391 AX601391 Sequence
19	107	100.0	2662	9	HSFGRREL X52833 Human flg m
20	107	100.0	2733	6	AR019675 AR019675 Sequence
21	107	100.0	2733	9	HSFGR X51803 Human mRNA
22	107	100.0	2877	10	BC033447 BC033447 Mus muscu
23	107	100.0	2912	10	AF176552 AF176552 Mus muscu
24	107	100.0	2983	10	MUSFGFR M33760 Mouse FGF r
25	107	100.0	3198	9	HUMHBGFA M63887 Human hepar
26	107	100.0	3223	9	HUMHBGFBV E03335 Human b-FGF
27	107	100.0	3328	6	E03335 AX774782 Sequence
28	107	100.0	3328	6	AX774782 M37722 Human short
29	107	100.0	3328	9	HUMBFGFS AX658247 Sequence
30	107	100.0	3343	6	AX658247 M34641 Human fibro
31	107	100.0	3343	9	HUMFGF1A M34185 Human fibro
32	107	100.0	3365	9	HUMFGF2H BC053921 Mus muscu
33	107	100.0	3419	10	BC053921 AR020617 Sequence
34	107	100.0	3503	6	AR020617 M34186 Human fibro
35	107	100.0	3503	9	HUMFGF3H E03799 BC010200 Mus muscu
36	107	100.0	3525	6	E03799 M60485 Human fibro
37	107	100.0	3850	10	BC010200 X66945 H.sapiens N
38	107	100.0	3901	9	HUMFGFAA AX587712 Sequence
39	107	100.0	3981	9	HSNSAMTK BC018128 Homo sapi
40	107	100.0	4066	6	AX587712 BX484967 Homo sapi
41	107	100.0	4086	9	BC018128 AX883990 Sequence
42	107	100.0	5543	9	AX883990 BD160651 Primer fo
43	101	94.4	5219	6	AX883990 AK024388 Homo sapi
44	101	94.4	5219	6	BD160651
45	101	94.4	5219	9	AK024388

ALIGNMENTS

RESULT 1

HUMHBGFC 1508 bp mRNA linear PRI 27-APR-1993
 LOCUS Human heparin-binding growth factor receptor (HBGF-R-alpha-a3)
 DEFINITION mRNA, complete cds.
 ACCESSION M63889
 VERSION M63889.1 GI:183882
 KEYWORDS heparin-binding growth factor receptor.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 1608)
 AUTHORS Hou, J.-Z., Kan, M.K., McKeehan, K., McBride, G., Adams, P. and McKeehan, W.L.
 TITLE Fibroblast growth factor receptors from liver vary in three structural domains
 JOURNAL Science 251 (4994), 665-668 (1991)
 MEDLINE 91126480
 PUBMED 1846977
 COMMENT Original source text: Homo sapiens liver cDNA to mRNA.
 FEATURES
 Location/Qualifiers
 source
 1..1608
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /cell_type="hepatocyte"
 /tissue_type="liver"
 1..1608
 /gene="HBGF-R"
 /gene="HBGF-R"
 /codon_start=1
 /product="heparin-binding growth factor receptor"
 /protein_id="AAA35960.1"
 /db_xref="GI:183883"
 /translation="MWSKSLFWVLVATLCTARPSPLPQACPDLEAKSCSAS
 FHSITPLPFGILKLS"
 mat_peptide 41..1866
 /gene="flg"
 /product="tyrosine kinase"
 ORIGIN
 Alignment Scores:
 Pred. No.: 4,06e-09 Length: 1608
 Score: 107.00 Matches: 22
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 9 Gaps: 0
 US-09-757-415A-3 (1-22) x HUMHBGFC (1-1608)
 Qy 1 HisSerGlnMetAlaValHisLysLeuAlaLysSerIleProLeuArgArgGlnValThr 20
 Db 1102 CACAGCAGATGGCTGTGTCACAGCTGCCAAGAGCATCCCTCTGCCGACAGGTAACA 1161
 Qy 21 ValSer 22
 Db 1162 GTGTCT 1167
 RESULT 2
 HSFPLGMR 1974 bp mRNA linear PRI 28-JUL-1995
 LOCUS Human flg (fms-like gene) mRNA for putative protein tyrosine kinase (partial).
 DEFINITION
 ACCESSION Y00665
 VERSION Y00665.1 GI:558583
 KEYWORDS tyrosine kinase.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 1974)
 AUTHORS Ruta, M., Hawk, R., Ricca, G., Drohan, W., Zabelshansky, M., Laureys, G., Barton, D.E., Francke, U., Schlessinger, J. and Gliori, D.
 TITLE A novel protein tyrosine kinase gene whose expression is modulated

during endothelial cell differentiation
 Oncogene 3, 9-15 (1988)
 COMMENT On Oct 18, 1994 this sequence version replaced gi:31427.
 FEATURES
 Location/Qualifiers
 source
 1..1974
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 1..1971
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /chromosome="8p12"
 /clone="5"
 /cell_type="endothelial"
 /clone_lib="lambda-gtl1"
 1..1869
 /gene="flg"
 1..1869
 /gene="flg"
 /codon_start=1
 /product="tyrosine kinase"
 /protein_id="CAA68673.1"
 /db_xref="GI:558584"
 /translation="HRIIGYKRVATWSIIMDSVPSDKNYTCTIVENEYGSINHTYQ
 LDVVERSPHRIQAGLPANKTVALGNSVFMCXVSDPQPHIOMLKHIVNGSKI
 DNLPVQILKTAGVNTDKEMVHLNVSFEDAGEYTCLAGNSIGLSHSAWLTVE
 ALEORPAVMTSPLYLEIIYCTGAFLLSCVMGVSIVYKMSGTTKSPHQSMAVHLA
 KSIPLRQVTVSADSSASMSGVLVSPRLSSSGTSMLAGVSVEIPEPRWELPRD
 RIVLCKPLGEGCFQGVLAELGLDCKPNRTKAVAVMLKSDATEKDLSDISEMEM
 MMIGKKNIIINLIGACTQDGLYIVIVYASKNLEIYLQARRPPGLEICINPSHNE
 EQLSKDLVSCAYQVARGMEYLAASKCIHRDLAARNVLTEDNMKTLADFLGARDIHH
 IDYKKTNGRLPVKMAPEALFDRIYTHQSDVMSFGVLLWEIETLGGSPYGPVVEE
 FLKLLKEGRMDKPSNCTNELYMMWRDCHAVPSQRTFKQLVEDLDRIVALTISNQY
 LDLSNLDQVSPSPDTRSSCTSSGSDSVFSHEPLPEPCLPRHPAQLANGGLKRR"
 41..1866
 /gene="flg"
 /product="tyrosine kinase"
 ORIGIN
 Alignment Scores:
 Pred. No.: 5,08e-09 Length: 1974
 Score: 107.00 Matches: 22
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 9 Gaps: 0
 US-09-757-415A-3 (1-22) x HSFPLGMR (1-1974)
 Qy 1 HisSerGlnMetAlaValHisLysLeuAlaLysSerIleProLeuArgArgGlnValThr 20
 Db 625 CACAGCAGATGGCTGTGTCACAGCTGCCAAGAGCATCCCTCTGCCGACAGGTAACA 684
 Qy 21 ValSer 22
 Db 685 GTGTCT 690
 RESULT 3
 MMU23445 2259 bp mRNA linear ROD 06-APR-1995
 LOCUS Mus musculus fibroblast growth factor receptor-1, short isoform
 DEFINITION
 ACCESSION U23445
 VERSION U23445.1 GI:733537
 KEYWORDS
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 2259)
 AUTHORS Jin, Y., Pasumarthi, K.B., Bock, M.E., Lytras, A., Kardami, E. and

```

Cattini, P.A.
Cloning and expression of fibroblast growth factor receptor-1
isoforms in the mouse heart: evidence for isoform switching during
heart development
J. Mol. Cell. Cardiol. 26 (11), 1449-1459 (1994)
JOURNAL 95205422
MEDLINE 7897669
PUBMED
REFERENCE 2 (bases 1 to 2259)
AUTHORS Jin, Y.
TITLE Direct Submission
JOURNAL Submitted (24-MAR-1995) Yan Jin, Physiology, University of
Manitoba, 730 William Avenue, Winnipeg, Manitoba R3E 3J7, Canada
FEATURES
Location/Qualifiers
1..2259
/organism="Mus musculus"
/mol_type="mRNA"
/strain="Swiss Webster"
/db_xref="taxon:10090"
/clone="pBSmFGFR(S)"
/tissue_type="heart"
/dev_stage="embryo (15 days)"
58..2259
/note="FGFR-1(S); similar to mouse FGFR-1(L), Genbank
Accession Number U22324, but does not encode the first
immunoglobulin like domain of FGFR-1(L)"
CDS
/protein_id="AAC52183.1"
/db_xref="GI:733538"
/translation="MWGKCLLFWALVLTATLCTARPAPTLPEDQALPSSDDDDDD
SSSEKETDNTKPNRPVPVYWTSPERMEKKLHAPAAKTVKFKPSSGTNPPTLRWL
KNGFEKPDHRIIGYKYRYATWSIIMDSVPSDKGNYTCTIVENEYSGINHTYOLDVVE
RSPHRLPLAQLPANETVALGNSVEFMCKVYSDPQHLOWLKHLEVNGSKIGPNLRY
VQLTAGWTNTDMEVHLHNRVSPEDAGEYTCLAGNSGLSHSALWLVLEALER
PAVMTSPYLIIYCTGAFILSCMLGSVIIYKMGSKGSDHFSQMAVHKLAKSIPL
RQVTVSADSSASNGSVLLVPSRLSSGTPMAGSEVELEDPEDPWEIPLPRFLVIG
KPLGEGCGQGVLAELAEIGLDKPNRVTKAVKMLKSDATEKDLSDLSIEMEMKMTIG
KHKNIINLGLACTODGPLYTVIVASKNLREYLQARPPGLKEYCNPSPBEPQLSS
KDLVSCAYOVARGMEYLAKKCIHRDLAARNVLVDENVMKIDFGIARDIHHIDYVK
XTTNGRLPVKMAPEALPDRIYTHQSDVMSFGVLLWEIFLTGGSYPGVPEVEFKLL
KEGRMDKPSNCTNELYMMRDCWHAVPSQRPTEKQLVELDLDRIVALTSSQEYLDLSI
FLDQYSFSPFTSRSTCSGSDSVFSHEPPEEPCLPHPHTQLANSGLKRR"
mat_peptide
118..2256
/product="fibroblast growth factor receptor-1, short
isoform"
/note="FGFR-1(S)"

ORIGIN
Alignment Scores:
Pred. No.: 5,88e-09 Length: 2259
Score: 107.00 Matches: 22
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 10 Gaps: 0

US-09-757-415A-3 (1-22) x MMU23445 (1-2259)

Qy 1 HisSerGlnMetAlaValHisLysLeuAlaLysSerIleProLeuArgGlnValThr 20
|||||
Db 1015 CATGCCAGATGGCTGTGCACAGCTGCCAGAGCATCCCTCTGCGCAGACAGTAC 1074

Qy 21 ValSer 22
|||||
Db 1075 GTGTCA 1080

RESULT 4
AR380743
LOCUS AR380743 2360 bp DNA linear PAT 18-DEC-2003
DEFINITION Sequence 1288 from patent US 6607879.
ACCESSION AR380743
VERSION AR380743.1 GI:40088377

```

```

KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 2360)
AUTHORS Cocks, B.G., Stuart, S.G. and Seilhamer, J.J.
TITLE Compositions for the detection of blood cell and immunological
response gene expression
JOURNAL Patent: US 6607879-A 1288 19-AUG-2003;
FEATURES Location/Qualifiers
1..2360
/source
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Alignment Scores:
Pred. No.: 6.17e-09 Length: 2360
Score: 107.00 Matches: 22
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-09-757-415A-3 (1-22) x AR380743 (1-2360)

Qy 1 HisSerGlnMetAlaValHisLysLeuAlaLysSerIleProLeuArgGlnValThr 20
|||||
Db 1114 CACAGCCAGATGGCTGTGCACAGCTGCCAGAGCATCCCTCTGCGCAGACAGTAC 1173

Qy 21 ValSer 22
|||||
Db 1174 GTGTCT 1179

RESULT 5
HSFGR2IG
LOCUS HSFGR2IG 2360 bp mRNA linear PRI 12-SEP-1993
DEFINITION Human mRNA for fibroblast growth receptor 2-Ig domain.
ACCESSION X571122
VERSION X571122.1 GI:313366
KEYWORDS FGF receptor; fibroblast growth factor receptor; immunoglobulin
domain.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Bisemann, A., Ahn, J.A., Graziani, G., Tronick, S.R. and Ren, D.
TITLE Alternative splicing generates at least five different isoforms of
the human basic-FGF receptor
JOURNAL Oncogene 6 (7), 1195-1202 (1991)
MEDLINE 91319400
PUBMED 1650441
REFERENCE 2 (bases 1 to 2360)
AUTHORS Tronick, S.R.
TITLE Direct Submission
JOURNAL Submitted (02-JAN-1991) S.R. Tronick, NATIONAL INSTITUTES OF
HEALTH, NATIONAL CANCER INSTITUTE, BLDG 37/RM 1E24, BETHESDA
MARYLAND 20892, USA
COMMENT See also X57118-X57122.
FEATURES Location/Qualifiers
1..2360
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/cell_line="M432"
/cell_type="fibroblast"
/tissue_type="lung"
/dev_stage="embryo"
/germline
1..2360
/evidence="experimental"
163..2358
/codon_start=1

```


VHFDLQRLCRDRDVSINWLRDGVQVLAESNRTRITGEBVEVDQSPADSLGIYACV
 TSSPGSTDTTFSNVNADALPSSDDDDSSSEKEETDNTPNRPVAPYWTSPK
 MEKXDAVPAKTVKPCPSGNTPTLWLKNGKEPKPHRIGGYKVRVATWSIIMD
 SVVSDKNYTCIVENEYSINHLYQDNLVVERSHPRLQAGLPANKTVAGSNVGF
 CKVSDPQPHIOMLKHIVNGSKIGPNLYVOLLTAGNTTDEKNEVLHLRNVGF
 DAGETCLAGNSIGLSHSAWLVLEALERPAPWTSPLLEIIIICTGAPLISCMVG
 SVIVYKSGTGKSDFSQMAVHLKAKSIPURQVTVSADSSASMSGVLLVVRPSRLS
 SSGTFLMAGVSEYELPDRWELPDRVLVPLGPGEGFQGVLAELGDKDKPNRV
 TKVAVMLKSDATEKDLISSEMNMKVIKHKNIINLIGACTQDQGLVIVIVEYASK
 GNRLYVLAQRPGLLEYCNPSPHNPEQLSSKDLVSCAYOVARGMEYLAACKIHRDL
 AARNLVTDENVMTADFLGARDIHHIDYKKTNGRLPYKMMAPALPFRIVTHOSD
 VASFGVLLWEITFLGSGYPGVVPEELFKLKEGHRNDKPSNCTNELYMMRCDHNAV
 PSORTFQVLEDRIIVALTNOEYLDLNPLOYSFSPDPDRSSSTCSGSDSVFSH
 EPLPREPCLPRAQLANRGLKRR"

ORIGIN

Alignment Scores:
 Pred. No.: 6.48e-09 Length: 2469
 Score: 107.00 Matches: 22
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0

US-09-757-415A-3 (1-22) x A29216 (1-2469)

QY 1 HisserGlnMetAlaValHisLysLeuAlaLysSerIleProLeuArgGlnValThr 20
 DB 1225 CACAGCCAGATGGCTGTGCACAGCTGCCAAGAGCATCCCTCTGCGCAGACAGGTAACA 1284
 QY 21 ValSer 22
 DB 1285 GTGTCT 1290

RESULT 8
 I66313
 LOCUS 166313 2469 bp DNA linear PAT 28-DEC-1997
 DEFINITION Sequence 1 from patent US 5670323.
 ACCESSION 166313
 VERSION 166313.1 GI:2724290
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 Unclassified.
 REFERENCE 1 (bases 1 to 2469)
 AUTHORS Nova,M.Philip., Gonzalez,A.-M. and Baird,A.
 TITLE Process for the detection of malignant melanoma
 JOURNAL Patent: US 5670323-A 1 23-SEP-1997;
 FEATURES Location/Qualifiers
 source 1..2469
 /organism="unknown"
 /mol_type="unassigned DNA"

Alignment Scores:
 Pred. No.: 6.48e-09 Length: 2469
 Score: 107.00 Matches: 22
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0

ORIGIN

US-09-757-415A-3 (1-22) x I66313 (1-2469)
 QY 1 HisserGlnMetAlaValHisLysLeuAlaLysSerIleProLeuArgGlnValThr 20
 DB 1225 CACAGCCAGATGGCTGTGCACAGCTGCCAAGAGCATCCCTCTGCGCAGACAGGTAACA 1284
 QY 21 ValSer 22
 DB 1285 GTGTCT 1290

RESULT 9

AX481457
 LOCUS AX481457 2469 bp DNA linear PAT 16-AUG-2002
 DEFINITION Sequence 71 from Patent WO02055693.
 ACCESSION AX481457
 VERSION AX481457.1 GI:22316371
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1
 AUTHORS Kreutzer,R., Limmer,S., Rost,S. and Hadwiger,P.
 TITLE Method for inhibiting the expression of a target gene
 JOURNAL Patent: WO 02055693-A 71 18-JUL-2002;
 FEATURES Ribopharma AG (DE)
 Location/Qualifiers
 source 1..2469
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"

ORIGIN

Alignment Scores:
 Pred. No.: 6.48e-09 Length: 2469
 Score: 107.00 Matches: 22
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0

US-09-757-415A-3 (1-22) x AX481457 (1-2469)

QY 1 HisserGlnMetAlaValHisLysLeuAlaLysSerIleProLeuArgGlnValThr 20
 DB 1225 CACAGCCAGATGGCTGTGCACAGCTGCCAAGAGCATCCCTCTGCGCAGACAGGTAACA 1284
 QY 21 ValSer 22
 DB 1285 GTGTCT 1290

RESULT 10

RATFGFR1
 LOCUS RATFGFR1 2469 bp mRNA linear ROD 01-FEB-2000
 DEFINITION Rat mRNA for FGF receptor-1, complete cds.
 ACCESSION D12498
 VERSION D12498.1 GI:220737
 KEYWORDS FGF receptor-1; fibroblast growth factor receptor.
 SOURCE Rattus norvegicus (Norway rat)
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus
 REFERENCE 1 (bases 1 to 2469)
 AUTHORS Yazaki,N., Hiroko,F., Mitsuhiro,O., Toshisuke,K. and Nobuyuki,I.
 TITLE The structure and expression of the FGF receptor-1 mRNA isoforms in
 rat tissues
 JOURNAL Biochim. Biophys. Acta 1172, 37-42 (1993)
 REFERENCE 2 (bases 1 to 2469)
 AUTHORS Nobuyuki,I.
 TITLE Direct Submission
 JOURNAL Submitted (26-JUN-1992) Itoh Nobuyuki, Kyoto University Faculty of
 Pharmaceutical Sciences, Department of Biological Chemistry;
 Sakyo-ku, Kyoto, Kyoto 606-01, Japan (Tel:81-75-753-4562,
 Fax:81-75-761-8949)
 COMMENT Submitted (26-JUN-1992) on tape to DDBJ by:
 Nobuyuki Itoh
 Department of Biological Chemistry
 University Faculty of Pharmaceutical Sciences Sakyo-ku, Kyoto
 606-01
 Japan
 Phone: 075-753-4562
 Fax: 075-761-8949
 Location/Qualifiers

FEATURES

```

source
1. .2469
/organism="Rattus norvegicus"
/mol_type="mRNA"
/strain="Wistar"
/db_xref="taxon:10116"
/sex="male"
/tissue_type="brain"
/dev_stage="adult"
1. .2469
/codon_start=1
/product="FGF receptor-1"
/protein_id="BAA02059.1"
/db_xref="GI:220738"
/translation="MWGRGLLFWAVLVTALCTARPAPTLPQAOQWGVVEVESLL
VHFGDQLQLRDLDDQVSNLWLDGVLQESNRITITGEEVEVDSDIPADSGLYACV
TNSPSSDTTFVSNVSDALPSSDDDDDDSSSEKEETDNTKPNRPVAPYMTSIPK
MEKHLHAPAAKTVKFCPSGTPPTLWMLKNGKEPKPDHRIIGYKRVATWSIIND
SVYSDKNGYTCIENEYGSINHLYQDLDVSPHRLPQAGLANKTVALGSNVEM
CKVSDPOPHIOMLKHIEVNGSKIGPDNLPIYDQILKTAGVNTTKEMEVLHNRVSE
DAGEYTCLAGNSIGLSHSAMTVEALEERPAVMTSPLYLEIIITCTGAFLISCMVG
SVLITKMSGKISDFHSCQMAVHLAKSIPLRRQVTVSADSSASMSGVLIVRPSRL
SVCTPMLAGVSEYELFEDPRWELPDRVLVGLKPLGEGCFQGVLAELGDKDKPNRV
TKVAYMLKSDATEKDLSDLISEMENMMKMGKHNIINLLCACTQDGLPVIVVEYASK
GNLREYLQARRPPGLECYNPSHNPSEOLSKDLVSCAYQVARGMEYLAASKCIHRL
AARNVLTVEDNMKIADFGIARDIHIIDYKKTNGSLPVKMAPEALFDRITYHQSD
VMSQVLWELFTLGGSPNPGVPVVEELFKLLKEGHRMDKSCNTNELYMMRDCWNAV
PSQRTFQLVEDLDRIVALTNSQEYLDLSIPLOQYSPSPFDRSTRSCSSGSDSVFSH
BFFPEPCPLRPHPTQLANGGLNRR"
1. .60
sig_peptide
mat_peptide
1. .2466
/product="FGF receptor-1"

ORIGIN
Alignment Scores:
Pred. No.: 6.48e-09 Length: 2469
Score: 107.00 Matches: 22
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 10 Gaps: 0

US-09-757-415A-3 (1-22) x RATFGFR1 (1-2469)
Qy 1 HisSerGlnMetAlaValHisLysLeuAlaLysSerIleProLeuArgArgGlnValThr 20
Db 1225 CATAGCAGATGCTGTGCTAGTACGCTAGAGCATCCCTCTCCGACAGGTAACA 1284
Qy 21 ValSer 22
Db 1285 GTGTCA 1290

RESULT 11
MUSFGFBC 2499 bp mRNA linear ROD 27-APR-1993
LOCUS Mouse fibroblast growth factor mRNA, complete cds.
DEFINITION
ACCESSION M65053
VERSION M65053.1 GI:193292
KEYWORDS fibroblast growth factor.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
1 (bases 1 to 2499)
Kouhara.H., Kasayama.S., Saito.H., Matsumoto.K. and Sato.B.
Expression cDNA cloning of fibroblast growth factor (FGF) receptor
in mouse breast cancer cells: a variant form in FGF-responsive
transformed cells
JOURNAL Biochem. Biophys. Res. Commun. 176 (1), 31-37 (1991)
MEDLINE 91207411
PUBMED 1708247
ORIGINAL SOURCE text: Mouse FGF-responsive breast cancer cell line
SC-3, cDNA to mRNA.
FEATURES
Location/Qualifiers

```

```

source
1. .2499
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/cell_line="FGF-responsiveSC-3"
/tissue_type="breast cancer"
1. .2499
/genes="FGF"
1. .2499
/genes="FGF"
/codon_start=1
/product="fibroblast growth factor"
/protein_id="AAA37620.1"
/db_xref="GI:193293"
/translation="MWGWKCLLFWAVLVTALCTARPAPTLPQSSSSWPLWVAAAAQ
PWGVFVSVSLVHFGDQLQLRDLDDQVSNLWLDGVLQESNRITITGEEVEVD
SIPADSGLYACVTSFSSGSDTFVSNVSDALPSSDDDDSSSEKEETDNTKPNRP
VAPYMTSPKMEKHLHAPAAKTVKFCPSGTPPTLWMLKNGKEPKPDHRIIGYKRV
ATWSIINDSVDPDKNGYTCIENEYGSINHLYQDLDVSPHRLPQAGLANKTVALGSN
VEMCKVSDPOPHIOMLKHIEVNGSKIGPDNLPIYDQILKTAGVNTTKEMEVLHNRV
SEDAEYTCLAGNSIGLSHSAMTVEALEERPAVMTSPLYLEIIITCTGAFLISCMVG
SVLITKMSGKISDFHSCQMAVHLAKSIPLRRQVTVSADSSASMSGVLIVRPSRL
SVCTPMLAGVSEYELFEDPRWELPDRVLVGLKPLGEGCFQGVLAELGDKDKPNRV
TKVAYMLKSDATEKDLSDLISEMENMMKMGKHNIINLLCACTQDGLPVIVVEYASK
GNLREYLQARRPPGLECYNPSHNPSEOLSKDLVSCAYQVARGMEYLAASKCIHRL
AARNVLTVEDNMKIADFGIARDIHIIDYKKTNGSLPVKMAPEALFDRITYHQSD
VMSQVLWELFTLGGSPNPGVPVVEELFKLLKEGHRMDKSCNTNELYMMRDCWNAV
PSQRTFQLVEDLDRIVALTNSQEYLDLSIPLOQYSPSPFDRSTRSCSSGSDSVFSH
BFFPEPCPLRPHPTQLANGGLNRR"
1. .2499
sig_peptide
mat_peptide
1. .2499
/product="FGF receptor-1"

ORIGIN
Alignment Scores:
Pred. No.: 6.57e-09 Length: 2499
Score: 107.00 Matches: 22
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 10 Gaps: 0

US-09-757-415A-3 (1-22) x MUSFGFBC (1-2499)
Qy 1 HisSerGlnMetAlaValHisLysLeuAlaLysSerIleProLeuArgArgGlnValThr 20
Db 1255 CATAGCAGATGCTGTGCTAGTACGCTAGAGCATCCCTCTCCGACAGGTAACA 1314
Qy 21 ValSer 22
Db 1315 GTGTCA 1320

RESULT 12
B0016930 2526 bp RNA linear PAT 27-AUG-2002
LOCUS Antisense nucleic acid compound.
DEFINITION
ACCESSION B0016930
VERSION B0016930.1 GI:22558106
KEYWORDS JP 2001252085-A/5.
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1 (bases 1 to 2526)
AUTHORS Matsuda,Y., Tanaka,Y., Takeuchi,H. and Uchida,K.
TITLE Antisense nucleic acid compound
JOURNAL TOA GOSSEI CO LTD
COMMENT PN JP 2001252085-A/5
PD 18-SEP-2001
PI 15-MAR-2000 JP 2000072741
C12N15/09,A61K31/711,A61K48/00,A61P43/00,C12N15/00 CC
Strandedness: Single;
CC Topology: Linear;
FH Key Location/Qualifiers.
FEATURES
Location/Qualifiers

```

```

source
1. .2526
/organism="unidentified"
/mol_type="genomic RNA"
/db_xref="taxon:32644"

ORIGIN

Alignment Scores:
Pred. No.: 6,64e-09 Length: 2526
Score: 107.00 Matches: 22
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-09-757-415A-3 (1-22) x BD016930 (1-2526)

QY 1 HisSerGlnMetAlaValHisLysLeuAlaLysSerIleProLeuArgGlnValThr 20
Db 1282 CATAGCCAGATGGCTGTGCACAAAGCTGCCAAGAGCATCCCTCTGCCGACACAGGTAA 1341

QY 21 ValSer 22
Db 1342 GTGTCA 1347

RESULT 13
MMU22324
LOCUS
DEFINITION
Mus musculus fibroblast growth factor receptor-1 mRNA, long isoform
precursor, complete cds.
ACCESSION
U22324
VERSION
U22324.1 GI:722340
KEYWORDS
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
Jin,Y., Pasumarthi,K.B., Bock,M.E., Lytras,A., Kardami,E. and
Cattini,P.A.
Cloning and expression of fibroblast growth factor receptor-1
isoforms in the mouse heart: evidence for isoform switching during
heart development
J. Mol. Cell. Cardiol. 26 (11), 1449-1459 (1994)
95205422
PUBMED
7897669
REFERENCE
2 (bases 1 to 2526)
AUTHORS
Jin,Y.
Direct Submission
Submitted (06-MAR-1995) Yan Jin, Physiology, University of
Manitoba, 730 William Avenue, Winnipeg, Manitoba R3E 3J7, Canada
LOCATION/Qualifiers
1. .2526
/organism="Mus musculus"
/mol_type="mRNA"
/strain="Swiss Webster"
/db_xref="taxon:10090"
/clone="pBSmfGFR(L)"
/tissue_type="heart"
/dev_stage="embryo (15 days)"
58. .2526
/note="FGFR1(L); similar to mouse FGFR-1 encoded by
GenBank Accession Number M28998 and to human FGFR-1
encoded by GenBank Accession Numbers M34187 and M34188"
/codon_start=1
/product="fibroblast growth factor receptor-1, long
isoform precursor"
/protein_id="AAC52182.1"
/db_xref="GI:722341"
/translation="MWGKKLLFWAVLVATLCTARAPTLPQCAQWGPVVEVSELL
VSPGGLQRLRDVGSINWLVGSNTRITGEEVRSIFPADSGLYACV
TSPGSGDTTYFNVNVDALPSSDDDDDDSSKEKETDNTKPNRPVAPYWTISPEK
MEKHLHVAAPKAPKCPSSGNTPLRLWLNKGKFKPDHRIKGVKRVATWISIMD
SVVPSDKGNYTCIVENEYGSINHTYQLDVVERSPHPILOAGLPANETVALGNSVFEM

CDS
/note="FGFR1(L); similar to mouse FGFR-1 encoded by
GenBank Accession Number M28998 and to human FGFR-1
encoded by GenBank Accession Numbers M34187 and M34188"
/codon_start=1
/product="fibroblast growth factor receptor-1, long
isoform precursor"
/protein_id="AAC52182.1"
/db_xref="GI:722341"
/translation="MWGKKLLFWAVLVATLCTARAPTLPQCAQWGPVVEVSELL
VSPGGLQRLRDVGSINWLVGSNTRITGEEVRSIFPADSGLYACV
TSPGSGDTTYFNVNVDALPSSDDDDDDSSKEKETDNTKPNRPVAPYWTISPEK
MEKHLHVAAPKAPKCPSSGNTPLRLWLNKGKFKPDHRIKGVKRVATWISIMD
SVVPSDKGNYTCIVENEYGSINHTYQLDVVERSPHPILOAGLPANETVALGNSVFEM

```

```

/db_xref="GI:309116"
/translation="MWGKCLLFWALVLTATLCTARPAFTLPEQAPGWVPEVESLL
VHPGLLQRLRDVQSNLWRDGVQVLESNRTITGEEVEVRDIPADSLGYACV
TSSPGSDTTTFYFNVNVDALPSSDDDDDDSSSEKETDNTKPNRPVAPYWTSPK
MEKLAHVAATKVKFKPCSSGTPNPTLRLWLNKGKFKPDHRIQGYKRVATWSIID
KVYSDPKNCTYICVENEYGSINHLYQDVVERSPHPSAGLPANKTVAGSNVREM
SWYSDPQPHQIOWLKHIEVNGSKIGPNLPYQILKTAGVNTTDKENEVLHLNVEFE
DAGETCLAGNSIGLSHSAWLTVEALEERPAVNTSPLEIIITCTGAFLSCMLG
SVIYKMGSGTKSDPFSQMAVHLKASIPLRQVTVSADSSASMSGVLVLRPSRLS
SSGTFAGVSEYELPEDRWELPRLVLGKPLGEGCFQGVLAELAGLDKDKPNRV
TKVAVKMLSDATEKDLISSEMMMKMGKHNIINLLGACTQDGLYVIIVEYASK
GNLREYLARPPGLGYCNPSPNEQLSKDLYSCAYOVARGMEYLAKCKIHRDL
AARVLTEDNVMTADGLARDIHHIDYKKTNGRLPVKMWMAPEALFRIYTHOSD
VWSGVLLWEFTLGGSPYGPVVEELFKLLKEGRMDKPSNCTNELYMMRDCWAVPS
PSQPTFKQLVEDLDRIVALTNSQEYLDLSIPLDQYSPFPDTRSSCTSSGSDSVFSE
EPLPEPCLPHPHTQLANSGLKRR"
58.117
/note="fibroblast growth factor-receptor signal peptide
(put.); putative"
118.2523
/product="fibroblast growth factor-receptor"

sig_peptide
mat_peptide

ORIGIN

Alignment Scores:
Pred. No.: 6.64e-09 Length: 2526
Score: 107.00 Matches: 22
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 10 Gaps: 0

US-09-757-415A-3 (1-22) x MUSBFGR (1-2526)

QY 1 HisSerGlnMetAlaValHisLysLeuAlaLysSerIleProLeuArgGlnValThr 20
Db 1282 CATAGCCAGATGGCTGTGCACAAAGTGGCAAGATCCCTCTGCGGACAGGTAACA 1341

QY 21 ValSer 22
Db 1342 GTGTCA 1347

RESULT 15
MMFGF
LOCUS
DEFINITION
Mouse fms-like gene (mflg) mRNA for fibroblast growth factor
receptor (MFR).
ACCESSION
X51893
VERSION
X51893.1 GI:50959
KEYWORDS
fibroblast growth factor receptor; growth factor receptor; kinase;
proto-oncogene; receptor; transmembrane protein; tyrosine kinase.
SOURCE
Mus musculus (house mouse)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 2625)
Safran,A., Avivi,A., Orr-Urtreger,A., Neufeld,G., Lonai,P.,
Givol,D. and Yarden,Y.
The murine fig gene encodes a receptor for fibroblast growth factor
Oncogene 5 (5), 635-643 (1990)
MEDLINE
90265603
PUBMED
2161096
REFERENCE
2 (bases 1 to 2625)
Safran,A.
Direct Submission
TITLE
Submitted (20-FEB-1990) Safran A., The Weizmann Institute of
Science, Rehovot 76100, Israel
JOURNAL
Science, Location/Qualifiers
FEATURES
source
1..2625
/organism="Mus musculus"
/mol_type="mRNA"
/strain="BALB/c"
/db_xref="taxon:10090"
/tissue_type="brain"

```

```

/clone_lib="lambda gtl1"
121..2583
/codon_start=1
/product="precursor polypeptide (AA -21 to 799)"
/protein_id="CAA36175.1"
/db_xref="GI:50960"
/db_xref="GOA:P16092"
/db_xref="SWISS-PROT:P16092"
/translation="MWGKCLLFWALVLTATLCTARPAFTLPEQAPGWVPEVESLL
VHPGLLQRLRDVQSNLWRDGVQVLESNRTITGEEVEVRDIPADSLGYACV
TSSPGSDTTTFYFNVNVDALPSSDDDDDDSSSEKETDNTKPNRPVAPYWTSPKME
KLAHVAATKVKFKPCSSGTPNPTLRLWLNKGKFKPDHRIQGYKRVATWSIIDV
VYSDPQPHQIOWLKHIEVNGSKIGPNLPYQILKTAGVNTTDKENEVLHLNVEFE
GYTCLAGNSIGLSHSAWLTVEALEERPAVNTSPLEIIITCTGAFLSCMLGSSV
IYKMGSGTKSDPFSQMAVHLKASIPLRQVTVSADSSASMSGVLVLRPSRLSS
SVIYKMGSGTKSDPFSQMAVHLKASIPLRQVTVSADSSASMSGVLVLRPSRLS
SSGTFAGVSEYELPEDRWELPRLVLGKPLGEGCFQGVLAELAGLDKDKPNRV
TKVAVKMLSDATEKDLISSEMMMKMGKHNIINLLGACTQDGLYVIIVEYASKGN
GNLREYLARPPGLGYCNPSPNEQLSKDLYSCAYOVARGMEYLAKCKIHRDLAA
AARVLTEDNVMTADGLARDIHHIDYKKTNGRLPVKMWMAPEALFRIYTHOSD
VWSGVLLWEFTLGGSPYGPVVEELFKLLKEGRMDKPSNCTNELYMMRDCWAVPS
PSQPTFKQLVEDLDRIVALTNSQEYLDLSIPLDQYSPFPDTRSSCTSSGSDSVFSE
EPLPEPCLPHPHTQLANSGLKRR"
121..183
/note="signal peptide (AA -21 to -1)"
184..2580
/product="fibroblast growth factor receptor (AA 1 to 799)"

ORIGIN

Alignment Scores:
Pred. No.: 6.93e-09 Length: 2625
Score: 107.00 Matches: 22
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 10 Gaps: 0

US-09-757-415A-3 (1-22) x MMFGF (1-2625)

QY 1 HisSerGlnMetAlaValHisLysLeuAlaLysSerIleProLeuArgGlnValThr 20
Db 1339 CATAGCCAGATGGCTGTGCACAAAGTGGCAAGATCCCTCTGCGGACAGGTAACA 1398

QY 21 ValSer 22
Db 1399 GTGTCA 1404

Search completed: August 8, 2004, 23:07:26
Job time : 814.618 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 8, 2004, 19:43:05 ; Search time 86.8421 Seconds
(without alignments)
1076.210 Million cell updates/sec

Title: US-09-757-415A-3

Perfect score: 107

Sequence: 1 HSCMAVHKLAKSLPLRRQVTVS 22

Scoring table:

BLOSUM62	Xgapop 10.0	Xgapext 0.5
	Ygapop 10.0	Ygapext 0.5
	Fgapop 6.0	Fgapext 7.0
	Delop 6.0	Delext 7.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh
-Q=/cgn2_1/USPIO_spool/US09757415/runat_02082004_101230_6675/app_query.fasta_1.526
-DB=N Geneseg 29Jan04 -Qfix=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0
-LOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFWT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USBR=US09757415 @CGN 1.1 596 @runat_02082004_101230_6675 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEX=7

Database : N Geneseg 29Jan04: *
1: Genesegn1980s: *
2: Genesegn1990s: *
3: Genesegn2000s: *
4: Genesegn2001as: *
5: Genesegn2001bs: *
6: Genesegn2002s: *
7: Genesegn2003as: *
8: Genesegn2003bs: *
9: Genesegn2003cs: *
10: Genesegn2004s: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	107	100.0	2469	2 AAQ13311	AAQ13311 Basic FGF
2	107	100.0	2469	6 ABV78187	ABV78187 Human FGF
3	107	100.0	2469	6 ABZ35763	ABZ35763 Human FGF
4	107	100.0	2469	6 ABX10006	ABX10006 Human FGF
5	107	100.0	2469	6 ABL91728	ABL91728 Human pol
6	107	100.0	2526	6 ABA01689	ABA01689 Mouse fib
7	107	100.0	2662	2 AAQ21003	AAQ21003 flg recep
8	107	100.0	2662	7 AAD52780	AAD52780 Human FGF

9	107	100.0	2733	2 AAT31051	AAT31051 Human fib
10	107	100.0	2733	9 ACC42967	ACC42967 Human fib
11	107	100.0	2856	2 AAQ10448	AAQ10448 Human bas
12	107	100.0	3328	2 AAQ20914	AAQ20914 Human bFG
13	107	100.0	3328	6 ABZ35187	ABZ35187 Human gen
14	107	100.0	3328	9 ADEB4879	ADEB4879 Farnesyl
15	107	100.0	3343	2 AAX87179	AAX87179 Fibroblas
16	107	100.0	3454	7 ABX63048	ABX63048 Human cDN
17	107	100.0	3503	2 AAV44041	AAV44041 Mouse bFG
18	107	100.0	3525	2 AAQ27658	AAQ27658 N-sam cDN
19	107	100.0	3726	3 AAF18252	AAF18252 Lung canc
20	107	100.0	4066	6 ABV94191	ABV94191 Breast ca
21	107	100.0	4628	6 ABS65210	ABS65210 cDNA enco
22	101	94.4	5219	4 AAH18659	AAH18659 Human cDN
23	101	94.4	11294	4 AAL04906	AAL04906 Human rep
24	101	94.4	11294	4 ABL97800	ABL97800 Human tes
25	101	94.4	15123	4 AAL04907	AAL04907 Human rep
26	101	94.4	15123	4 ABL97801	ABL97801 Human tes
27	87	81.3	505	8 ACH17485	ACH17485 Human adu
28	84	78.5	1451	6 ABK63559	ABK63559 Rat seque
29	83	77.6	1954	2 AAQ14849	AAQ14849 Clone pTB
30	83	77.6	2345	2 AAQ14049	AAQ14049 Human act
31	83	77.6	2345	2 AAQ10868	AAQ10868 KGF recep
32	83	77.6	2466	6 ABV78201	ABV78201 Human RGF
33	83	77.6	2466	6 ABZ35777	ABZ35777 Human RGF
34	83	77.6	2466	6 ABX10020	ABX10020 Human FGF
35	83	77.6	2466	6 ABL91742	ABL91742 Human pol
36	83	77.6	2676	2 AAQ14850	AAQ14850 Clone pTB
37	83	77.6	3416	2 AAQ21004	AAQ21004 bek recep
38	83	77.6	4268	6 ABL64983	ABL64983 Lung canc
39	83	77.6	4268	6 ABL66289	ABL66289 Lung canc
40	83	77.6	4268	6 ABL68594	ABL68594 Kidney ca
41	81	75.7	2675	2 AAV19076	AAV19076 Fibroblas
42	81	75.7	2675	2 AAX85614	AAX85614 Bek-like
43	81	75.7	2681	2 AAV19078	AAV19078 Keratinoc
44	81	75.7	2681	2 AAX85616	AAX85616 Newt kera
45	80	74.8	99	7 ACF19037	ACF19037 Tumour ce

ALIGNMENTS

RESULT 1
AAQ13311
ID AAQ13311 standard; cDNA; 2469 BP.
XX
AC AAQ13311;
XX
XX
DT 25-MAR-2003 (revised)
DT 28-OCT-1991 (first entry)
XX
DE Basic FGF receptor.
XX
KW Basic fibroblast growth factor; human; ss.
XX
OS Homo sapiens.

Key	Location/Qualifiers
CDS	1..2469
FT	/tag= a
FT	sig_peptide 1..63
FT	/tag= b
FT	mat_peptide 64..2466
FT	/tag= c

XX	WO9111459-A.
PN	
XX	
PD	08-AUG-1991.
XX	
PF	23-JAN-1990; 90GB-00001466.
XX	
PR	23-JAN-1990; 90GB-00001466.
XX	
PA	(FARM) FARMITALIA ERBA SRL CARLO.

XX Bergonzoni L, Mazue G, Isacchi A, Roncucci R, Sarmientos P;
 XX WPI; 1991-252611/34.
 DR P-PSDB; AARI3549.
 XX Extracellular form of human fibroblast growth factor receptor - used to
 PT treat tumours, abnormal angiogenesis e.g. diabetic retinopathy,
 PT rheumatoid arthritis and arteriosclerosis and as contraceptives.
 XX Claim 3; Fig 2; 29pp; English.
 XX The sequence was obtd. from two overlapping clones, PL5 and PL10,
 CC isolated from a placentar lambda gt11 cDNA library. The DNA can be used
 CC to express recombinant bFGF receptor which is an antagonist of human aFGF
 CC and bFGF. The receptor can be used to treat abnormal angiogenesis (e.g.
 CC in diabetic retinopathy, neovascular glaucoma, etc, and possibly certain
 CC solid tumours), and in contraceptives. Dosage is 10-100 ug. See also
 CC AAQ13308-Q13311. (Updated on 25-MAR-2003 to correct PA field.)
 XX
 SQ Sequence 2469 BP; 576 A; 712 C; 698 G; 483 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 7.56e-09 Length: 2469
 Score: 107.00 Matches: 22
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 2 Gaps: 0

US-09-757-415A-3 (1-22) x AAQ13311 (1-2469)

QY 1 HisSerGlnMetAlaValHisLysLeuAlaLysSerIleProLeuArgArgGlnValThr 20
 Db 1225 CACAGCCAGATGGTGTGCACAAAGCTGGCCAGAGCATCCCTCTGCGCAGACAGTAACA 1284
 QY 21 ValSer 22
 Db 1285 GTGCT 1290

RESULT 2
 ABV78187
 ID ABV78187 standard; DNA; 2469 BP.
 XX
 AC ABV78187;
 XX
 DT 15-NOV-2002 (first entry)
 XX
 DE Human FGFR1 DNA SEQ ID NO 71.
 XX
 KW RNA inhibition; dsRNA; gene expression inhibitor; oncogene; cytostatic;
 KW virucide; protozoacide; gene; ds.
 XX
 OS Homo sapiens.
 XX
 PN WO200255693-A2.
 XX
 PD 18-JUL-2002.
 XX
 PF 09-JAN-2002; 2002WO-EP000152.
 XX
 PR 09-JAN-2001; 2001DE-01000586.
 PR 26-OCT-2001; 2001DE-01055280.
 PR 29-NOV-2001; 2001DE-01058411.
 PR 07-DEC-2001; 2001DE-01060151.
 XX
 PA (RIBO-) RIBOPHARMA AG.
 XX
 PI Kreutzer R, Limmer S, Rost S, Hadwiger P;
 XX
 DR WPI; 2002-590671/63.
 XX
 PT Inhibiting expression of target gene, useful e.g. for inhibiting

PT oncogenes, by administering double-stranded RNA complementary to the
 PT target and having an overhang.
 XX
 PS Claim 10; Page 150-151; 203pp; German.
 XX
 CC The invention relates to inhibiting expression of a target gene (I) in a
 CC cell by introducing an inhibitory RNA (dsRNA1) having a double-stranded
 CC structure of at most 49 consecutive bases. At least part of one strand
 CC (as1) of dsRNA1 is complementary to (I) and at least one end of dsRNA1
 CC has an overhang of 1-4 nucleotides. The method is used to inhibit the
 CC expression of a wide range of genes, e.g. oncogenes, cytokine genes etc.
 CC in humans, also genes in Plasmodium or in viruses or viroids that are
 CC pathogenic for humans, animals or plants. Introducing an overhang into
 CC dsRNA greatly increases effectiveness for inhibiting gene expression,
 CC both in vivo and in vitro and also increases stability and thus the
 CC effective concentration inside the cell. The present sequence is that of
 CC a gene related to the invention
 XX
 SQ Sequence 2469 BP; 576 A; 712 C; 698 G; 483 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 7.56e-09 Length: 2469
 Score: 107.00 Matches: 22
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0

US-09-757-415A-3 (1-22) x ABV78187 (1-2469)

QY 1 HisSerGlnMetAlaValHisLysLeuAlaLysSerIleProLeuArgArgGlnValThr 20
 Db 1225 CACAGCCAGATGGTGTGCACAAAGCTGGCCAGAGCATCCCTCTGCGCAGACAGTAACA 1284
 QY 21 ValSer 22
 Db 1285 GTGCT 1290

RESULT 3
 ABZ35763
 ID ABZ35763 standard; DNA; 2469 BP.
 XX
 AC ABZ35763;
 XX
 DT 07-FEB-2003 (first entry)
 XX
 DE Human FGFR1 polynucleotide SEQ ID NO 71.
 XX
 KW Double stranded RNA; dsRNA; RNAi; RNA inhibition; cytostatic; virucide;
 KW protozoacide; gene expression; antisense; tumour; infection; Plasmodium;
 KW virus; viroid; anti-GFP; human; HIV; human immunodeficiency virus;
 KW Hepatitis C virus; human papilloma virus; gene; ds.
 XX
 OS Homo sapiens.
 XX
 PN DE10100588-A1.
 XX
 PD 18-JUL-2002.
 XX
 PF 09-JAN-2001; 2001DE-01000588.
 XX
 PR 09-JAN-2001; 2001DE-01000588.
 XX
 PA (RIBO-) RIBOPHARMA AG.
 XX
 PI Kreutzer R, Limmer S, Rost S, Hadwiger P;
 XX
 DR WPI; 2002-683450/74.
 XX

PT Inhibiting expression of target genes, useful e.g. for treating tumors,
 PT by introducing into cells two double-stranded RNAs that are complementary
 PT to the target.
 XX

PS Claim 13; Page 47; 100pp; German.

XX The invention relates to inhibiting expression of a target gene in a cell

CC by introducing at least two oligoribonucleotides (dsRNAI and II), both

CC with a double-stranded (ds) structure of at most 49 sequential nucleotide

CC pairs. At least part of one strand (S1, S2) of the ds structures in each

CC of dsRNAI and II are complementary to regions in the target gene. The

CC method uses antisense inhibition of gene expression using double stranded

CC RNA inhibition (RNAi). The method is particularly used to treat tumours

CC or infections, especially by Plasmodium or viruses/viroids (pathogenic on

CC humans, animals or plants). The method provides more effective inhibition

CC of expression than known methods using a single dsRNA, even at very low

CC concentrations. When dsRNA has at least one unpaired nucleotide at the

CC end, stability (and thus effective concentration in the cell) is improved

CC and efficiency can be increased further by pretreating the cells with

CC interferon. The present sequence is that of a target DNA of the invention

XX

SQ Sequence 2469 BP; 576 A; 712 C; 698 G; 483 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 7.56e-09 Length: 2469

Score: 107.00 Matches: 22

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 100.00% Indels: 0

DB: 6 Gaps: 0

US-09-757-415A-3 (1-22) x ABZ35763 (1-2469)

QY 1 HisSerGlnMetAlaValHisLysLeuAlaLysSerIleProLeuArgArgGlnValThr 20

DB 1225 CACAGCCAGATGGCTGTGCACAGAGCTGGCCAGAGCATCCCTCTGGCAGACAGGTAACA 1284

QY 21 ValSer 22

DB 1285 GTGTCT 1290

RESULT 4

ABX10006

ID ABX10006 standard; DNA; 2469 BP.

XX

AC ABX10006;

XX

XX 23-JAN-2003 (first entry)

DT

DE Human FGFR1 DNA fragment SEQ ID 71.

DE

XX Oligoribonucleotide; interferon; oncogene; cytokine; Id; developmental;

XX prion; inhibition; human; ds.

XX

OS Homo sapiens.

XX

PN DE10100587-C1.

XX

XX 21-NOV-2002.

PD

XX

XX 09-JAN-2001; 2001DE-01000587.

PF

XX

XX 09-JAN-2001; 2001DE-01000587.

PR

XX

XX (RIBO-) RIBOPHARMA AG.

PA

XX

XX Kreutzer R, Limmer S, Rost S, Hadwiger P;

PI

XX

XX WPI; 2002-742209/81.

DR

XX Inhibiting expression of target genes, e.g. oncogenes, in cells, by

PT introduction of complementary double-stranded oligoribonucleotide, after

PT treating the cell with interferon.

PT

XX

XX Disclosure; Page 52; 98pp; German.

PS

XX This invention describes a novel method for inhibiting expression of a

CC

CC target gene by introducing into the cell that contains the target gene at

CC least one oligoribonucleotide (dsRNAI) that has a double-stranded (ds)

CC structure of not more than 49 consecutive nucleotides (nt), where at

CC least a segment of one strand of the ds structure is complementary with

CC the target gene and the cells are treated with interferon before

CC introduction of dsRNAI. The method is used to inhibit expression of

CC target genes, particularly oncogenes, cytokine genes, Id (not defined)

CC protein genes; developmental or prion genes, or genes expressed in

CC pathogenic organisms (particularly plasmidia) or in viruses or viroids

CC (pathogenic in humans, animals or plants). Treating the cells with

CC interferon greatly increases the extent to which dsRNA can inhibit

CC expression of the target genes, and the effect is even greater when dsRNA

CC are modified to increase their stability. ABX0936-ABX10075 represent

CC gene fragments used to illustrate the method of the invention

XX

SQ Sequence 2469 BP; 576 A; 712 C; 698 G; 483 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 7.56e-09 Length: 2469

Score: 107.00 Matches: 22

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 100.00% Indels: 0

DB: 6 Gaps: 0

US-09-757-415A-3 (1-22) x ABX10006 (1-2469)

QY 1 HisSerGlnMetAlaValHisLysLeuAlaLysSerIleProLeuArgArgGlnValThr 20

DB 1225 CACAGCCAGATGGCTGTGCACAGAGCTGGCCAGAGCATCCCTCTGGCAGACAGGTAACA 1284

QY 21 ValSer 22

DB 1285 GTGTCT 1290

RESULT 5

ABL91728

ID ABL91728 standard; DNA; 2469 BP.

XX

AC ABL91728;

XX

XX 28-MAY-2002 (first entry)

DT

XX Human polynucleotide SEQ ID NO 71.

DE

XX

XX Human; HIV; HCV; gene expression; oligoribonucleotide; tumour; pathogen;

XX Plasmodium; virus; viroid; cytokine; prion; antisense oligonucleotide;

XX cytostatic; virucide; protozoacide; antibacterial; ds.

XX

OS Homo sapiens.

XX

PN DE10100586-C1.

XX

XX 11-APR-2002.

PD

XX

XX 09-JAN-2001; 2001DE-01000586.

PF

XX

XX 09-JAN-2001; 2001DE-01000586.

PR

XX

XX (RIBO-) RIBOPHARMA AG.

PA

XX

XX Kreutzer R, Limmer S, Rost S, Hadwiger P;

PI

XX

XX WPI; 2002-270454/32.

DR

XX Inhibiting gene expression in cells, useful for e.g. treating tumors, by

PT introducing double-stranded complementary oligoRNA having unpaired

PT terminal bases.

PT

XX

XX Claim 13; Page 50; 104pp; German.

PS

XX The invention relates to a method for inhibiting expression of a target

CC gene (ABL91658-ABL91797) in a cell by introducing at least one

CC

CC oligoribonucleotide that has a double-stranded structure consisting of at
 CC most 49 sequential nucleotide pairs, with at least part of one strand
 CC complementary with the target gene and has at least one end a single-
 CC stranded segment of 1-4 nt. The method provides oligoribonucleotides for
 CC antisense inhibition of gene expression useful e.g. for treating tumours
 CC but the oligoribonucleotides may also be directed against genes present
 CC in pathogens (e.g. plasmodium or viruses/viroids, pathogenic on humans,
 CC animals or plants) or against cytokine, Id, developmental or prion genes.
 CC The method provides more effective inhibition of gene expression than use
 CC of known oligonucleotides, probably because the unpaired overhang
 CC increases stability and thus intracellular concentration
 XX
 SQ Sequence 2469 BP; 576 A; 712 C; 698 G; 483 T; 0 U; 0 Other;

Alignment Scores: 7.56e-09 Length: 2469
 Pred. No.: 107.00 Matches: 22
 Score: 107.00 Conservative: 0
 Percent Similarity: 100.00% Mismatches: 0
 Best Local Similarity: 100.00% Indels: 0
 Query Match: 100.00% Gaps: 0
 DB: 6

US-09-757-415A-3 (1-22) x ABU91728 (1-2469)

QY 1 HisSerGlnMetAlaValHisLysLeuAlaLysSerIleProLeuArgGlnValThr 20
 DB 1225 CACACCCAGATGGCTGTCCACAGCTGCCAAGAGCATCCCTCTGCCGACAGACGTAACA 1284

QY 21 ValSer 22

DB 1285 GTGTCT 1290

RESULT 6

ABA01689

ID ABA01689 standard; mRNA; 2526 BP.

AC ABA01689;

XX 08-FEB-2002 (first entry)

XX Mouse fibroblast growth factor receptor 1 mRNA SEQ ID NO:5.

XX Fibroblast growth factor receptor 1; FGFR1; antisense oligonucleotide;
 KW phosphorothioate; diagnosis; ss.

XX Mus sp.

PN JP2001252085-A.

XX 18-SEP-2001.

PF 15-MAR-2000; 2000JP-00072741.

PR 15-MAR-2000; 2000JP-00072741.

PA (TOAG) TOA GOSSEI CHEM IND LTD.

DR WPI; 2002-029661/04.

XX An antisense nucleic acid compound used as a diagnostic and therapeutic
 PT agent.

PS Example; Page 7-8; 10pp; Japanese.

XX The present invention describes an antisense nucleic acid compound (I),
 CC used as a diagnostic and therapeutic agent. (I) comprises: (i) an
 CC antisense nucleic acid compound (II) having the base sequence TTGGGTTT;
 CC (ii) an antisense nucleic acid compound (III) having the above base
 CC sequence in which the base number is expressed by a base sequence between
 CC 10 and 60; and (iii) an antisense nucleic acid molecule (IV) comprising
 CC one of the 3 sequences given in ABA01686 to ABA01689, all comprising of
 CC 20 nucleic acids, and that inhibit the expression of fibroblast growth
 CC factor receptor 1 (FGFR1). (I) is useful as a treating agent, a

CC diagnostic agent and a research reagent. The present sequence represents
 CC mouse FGFR1 mRNA, which is given in the exemplification of the present
 CC invention

SQ Sequence 2526 BP; 576 A; 710 C; 712 G; 0 T; 528 U; 0 Other;

Alignment Scores: 7.78e-09 Length: 2526
 Pred. No.: 107.00 Matches: 22
 Score: 107.00 Conservative: 0
 Percent Similarity: 100.00% Mismatches: 0
 Best Local Similarity: 100.00% Indels: 0
 Query Match: 100.00% Gaps: 0
 DB: 6

US-09-757-415A-3 (1-22) x ABA01689 (1-2526)

QY 1 HisSerGlnMetAlaValHisLysLeuAlaLysSerIleProLeuArgGlnValThr 20
 DB 1282 CAUAGCCAGUGGUGUGCACAAGCUGGCCAAGAGCAUCCUCCUGGCAGACAGGAACA 1341

QY 21 ValSer 22

DB 1342 GUGUCA 1347

RESULT 7

AAQ21003

ID AAQ21003 standard; cDNA; 2662 BP.

XX AC AAQ21003;

XX 25-MAR-2003 (revised)

DT 20-MAY-1992 (first entry)

XX flg receptor protein gene.

XX Fibroblast growth factor receptor; heparin binding proteins;

KW tyrosine kinase; fms-like gene; CSF-1; PDGF; ss.

XX Homo sapiens.

XX Key Location/Qualifiers

FT CDS 87..2546

FT /*tag= a

FT /product= "flg"

FT misc_feature 1105

FT /*tag= b

FT /note= "insertion; see comments"

FT 1173..1174

FT /*tag= c

FT /note= "deletion; see comments"

XX WO9200999-A.

XX 23-JAN-1992.

XX 06-JUL-1990; 90US-00549587.

XX 06-JUL-1990; 90US-00549587.

XX (ROR) RORER INT HOLDINGS INC.

XX Dionne CA, Crumley G, Jaye MC, Schlessing J;

XX WPI; 1992-056827/07.

DR P-PSDB; AAR20750.

XX New fibroblast growth factor receptor proteins - useful in treating gf-
 PT mediated conditions e.g. angiogenesis of tumours, mitogenic effects in
 PT psoriasis, arthritis.

XX Claim 10; Fig 7; 65pp; English.

XX The cDNA sequence was obt'd. using as template the partial flg known

CC sequence (Ruta et al., Oncogene, 3: 9-15). A full length flg cDNA clone
CC was obt'd. by repeated rounds of PCR using 8 primers (AAQ23610- 17), and
CC was cloned into pMJ30. NH 3T3 cells were cotransfected with a 1:20
CC mixture of pSV2neo and an flg expression vector contg. the flg probe
CC inserted into pMJ30 immediately downstream of the SV40 promoter and
CC cytomegalovirus enhancer. Clones of flg transfected cells have increased
CC binding for aFGF receptors, the best being clone Nflg26 which was used
CC for DNA sequencing. The sequence shown contains an apparent insertion
CC (see features) which disrupts the reading frame. This is restored,
CC however, by a deletion downstream in the sequence. No reference is made
CC by the authors to these discrepancies, which are therefore assumed to be
CC typing errors. Receptor proteins encoded by this cDNA sequence may be
CC used in pharmaceutical compns. to inhibit undesirable heparin-binding
CC growth factor mediated cellular responses or to inhibit the binding of an
CC opportunistic pathogen to human cells. Such undesirable responses may be
CC growth factor stimulated angiogene- sis and vascularisation of tumours,
CC mitogenic effects in psoriasis, arthritis, atherosclerosis and benign
CC prostatic hyper trophy. The derived flg protein may also be used for
CC screening drugs for treatment of such problems. See also AAQ21004.
CC (Updated on 25-MAR-2003 to correct PA field.)

XX SQ Sequence 2662 BP; 613 A; 787 C; 740 G; 522 T; 0 U; 0 Other;

Alignment Scores: Pred. No.: 8.31e-09 Length: 2662
Score: 107.00 Matches: 22
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 2 Gaps: 0

US-09-757-415A-3 (1-22) x AAQ21003 (1-2662)

Qy 1 HisSerGlnMetAlaValHisLysLeuAlaLysSerIleProLeuArgGlnValThr 20
Db 1312 CACAGCCAGATGGCTGTGCACAGCTGCCAAGAGCATCCCTCTGCCACAGAGTAA 1371
Qy 21 ValSer 22
Db 1372 GTGTCT 1377

RESULT 8
AAD52780
ID AAD52780 standard; DNA; 2662 BP.

XX AC AAD52780;
XX DT 14-MAY-2003 (first entry)
XX DE Human FGF-R1 DNA.

XX Human; pyridylpyrimidine derivative; cellular protein kinase; Scrapie;
KW cellular protein phosphatase; cellular signal transduction; prophylaxis;
KW prion infection; chronic wasting disease; CWD; Creutzfeldt-Jacob disease;
KW CJD; transmissible mink encephalopathy; bovine spongiform encephalopathy;
KW TME; BSE; Gerstmann-Strausler-Scheinker syndrome; GSS; Alpers syndrome;
KW fatal familial insomnia; FFI; kuru; neurodegenerative disease; neurotropic;
KW Alzheimer's disease; FGF-R1; flg; Flt-1; Flt-2; b-FGFR; gene; ds.

XX Homo sapiens.
XX FH Key Location/Qualifiers
XX CDS 88..2556
XX /*tag= a
XX /product= "Human FGF-R1 protein"

XX WO200293164-A2.
XX PD 21-NOV-2002.
XX PF 16-MAY-2002; 2002WO-EP005420.
XX PF 16-MAY-2001; 2001EP-00111858.

PR 29-MAY-2001; 2001US-0293528P.
PR 13-JUL-2001; 2001EP-00117113.
PR 18-JUL-2001; 2001US-0305898P.
XX (AXXI-) AXXIMA PHARM AG.
XX Stein-Gerlach M, Salassidis K, Bacher G, Mueller S;
XX WPI: 2003-120714/11.
XX P-PSDB; AAE34488.
XX New pyridylpyrimidine derivatives useful in the treatment or prevention
XX of infectious disease e.g. Kuru syndrome and Creutzfeld-Jacob disease
XX (CJD).
XX Disclosure; Page 61-62; 96pp; English.

XX The invention relates to novel pyridylpyrimidine derivatives and methods
XX of detecting prion infections and/or prion disease in an individual or in
XX cells, cell cultures and/or cell lysates. The method involves adding at
XX least one monoclonal or polyclonal antibody, oligonucleotide or pyridyl-
XX pyrimidine derivative to the sample or in cells, cell cultures and/or
XX cell lysates and detecting the activity of at least one human cellular
XX protein kinases (e.g.; FGF-R1 (also known as fig, Flt-1, Flt-2, b-FGFR),
XX Tkt (also known as CCK-2, DDR-2 or EDDR; EC number 2.7.1.112), Abl (also
XX known as c-abl), c-Kit, MKK7 (also known as SAPK1a, SAPKalpha), CDC2 (also
XX known as CDK1), PKA), human cellular protein phosphatases such as PTP-SL
XX (also known as MCP83) and PTP-zeta, the cellular signal transduction
XX molecules HSP80 and GPCR-1. The invention is useful for regulating the
XX production of prions in cells and in the manufacture of pharmaceutical
XX composition for prophylaxis and/or treatment of infectious disease (e.g.
XX Scrapie, chronic wasting disease (CWD), transmissible mink encephalopathy
XX (TME), Creutzfeldt-Jacob disease (CJD), bovine spongiform encephalopathy
XX (BSE), variant CJD, Gerstmann-Strausler-Scheinker syndrome (GSS), fatal
XX familial insomnia (FFI), Kuru and Alpers syndrome, especially BSE, CJD,
XX vCJD) or neurodegenerative diseases (e.g., Alzheimer's disease) in humans
XX or ruminants. The present sequence is human FGF-R1 DNA used in the
XX invention

SQ Sequence 2662 BP; 613 A; 787 C; 740 G; 522 T; 0 U; 0 Other;

Alignment Scores: Pred. No.: 8.31e-09 Length: 2662
Score: 107.00 Matches: 22
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 7 Gaps: 0

US-09-757-415A-3 (1-22) x AAD52780 (1-2662)

Qy 1 HisSerGlnMetAlaValHisLysLeuAlaLysSerIleProLeuArgGlnValThr 20
Db 1312 CACAGCCAGATGGCTGTGCACAGCTGCCAAGAGCATCCCTCTGCCACAGAGTAA 1371
Qy 21 ValSer 22
Db 1372 GTGTCT 1377

RESULT 9
AAT31051
ID AAT31051 standard; DNA; 2733 BP.

XX AC AAT31051;
XX DT 26-FEB-1997 (first entry)
XX DE Human fibroblast growth factor receptor 1 gene.

XX Human fibroblast growth factor receptor; FGFR1; alpha exon; pre-mRNA;
KW antisense oligonucleotide; inhibition; anti-tumour; glioma; glioblastoma;
KW ss.

```

OS XX Homo sapiens.
FH Key Location/Qualifiers
FT CDS 118..2586
FT /*tag= a
FT /*product= "FGFR1"
FT /*transl_except= (pos:2566..2568,aa:Gly)
XX WO2003038054-A2.
XX PN
XX PD 08-MAY-2003.
XX XX
XX PF 31-OCT-2002; 2002WO-US034986.
XX PR 31-OCT-2001; 2001US-0335583P.
XX PA (UYNV ) UNIV NEW YORK STATE.
XX PI Moosa M, Green DL, Linhard RJ;
XX DR WPI; 2003-482144/45.
XX DR P-PSDB; ABR56165.
XX PT Composition useful for identifying modulators of fibroblast growth factor
XX PT -mediated signaling for use in treating cancer, has ternary complex of
XX PT FGF-receptor, FGF ligand and heparin agonist or antagonist.
XX PS Disclosure; Fig 2B; 288pp; English.
XX CC The present invention relates to an isolated composition comprising a
XX CC ternary complex of an Fibroblast Growth Factor (FGF) ligand polypeptide
XX CC (ABR56163, ABR56165), an FGF receptor polypeptide (ABR56164) and a
XX CC heparin agonist or antagonist, where the agonist or antagonist binds to
XX CC the FGF ligand polypeptide and the FGF receptor polypeptide to form the
XX CC ternary complex. The composition is useful for identifying a compound
XX CC that is an inhibitor of FGF receptor activity. FGF1 is also known as
XX CC acidic FGF
XX SQ Sequence 2733 BP; 623 A; 810 C; 765 G; 532 T; 0 U; 2 Other;

Alignment Scores:
Pred. No.: 8.59e-09 Length: 2733
Score: 107.00 Matches: 22
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 2 Gaps: 0

US-09-757-415A-3 (1-22) x AAT31051 (1-2733)
QY 1 HisSerGlnMetAlaValHisIleuAlaLysSerIleProLeuArgGlnValThr 20
Db 1342 CACAGCCAGATGGCTGTGCACAAAGCTGCCAAGCATCCCTCTGCGCAGACAGGTAACA 1401
QY 21 ValSer 22
Db 1402 GTGTCT 1407

RESULT 10
ACC42967
ID ACC42967 standard; DNA; 2733 BP.
XX AC
XX ACC42967;
XX DT 18-DEC-2003 (first entry)
XX XX
XX DE Human Fibroblast Growth Factor Receptor 1 coding sequence.
XX KW Human; Fibroblast Growth Factor Receptor 1; FGFR1;
XX KW protein co-ordinate data; gene; ds.
XX OS Homo sapiens.

OS XX Homo sapiens.
FH Key Location/Qualifiers
FT CDS 118..2586
FT /*tag= a
FT /*product= "FGFR1"
FT /*transl_except= (pos:2566..2568,aa:Gly)
XX WO2003038054-A2.
XX PN
XX PD 08-MAY-2003.
XX XX
XX PF 31-OCT-2002; 2002WO-US034986.
XX PR 31-OCT-2001; 2001US-0335583P.
XX PA (UYNV ) UNIV NEW YORK STATE.
XX PI Moosa M, Green DL, Linhard RJ;
XX DR WPI; 2003-482144/45.
XX DR P-PSDB; ABR56165.
XX PT Composition useful for identifying modulators of fibroblast growth factor
XX PT -mediated signaling for use in treating cancer, has ternary complex of
XX PT FGF-receptor, FGF ligand and heparin agonist or antagonist.
XX PS Disclosure; Fig 2B; 288pp; English.
XX CC The present invention relates to an isolated composition comprising a
XX CC ternary complex of an Fibroblast Growth Factor (FGF) ligand polypeptide
XX CC (ABR56163, ABR56165), an FGF receptor polypeptide (ABR56164) and a
XX CC heparin agonist or antagonist, where the agonist or antagonist binds to
XX CC the FGF ligand polypeptide and the FGF receptor polypeptide to form the
XX CC ternary complex. The composition is useful for identifying a compound
XX CC that is an inhibitor of FGF receptor activity. FGF1 is also known as
XX CC acidic FGF
XX SQ Sequence 2733 BP; 623 A; 810 C; 765 G; 533 T; 0 U; 2 Other;

Alignment Scores:
Pred. No.: 8.59e-09 Length: 2733
Score: 107.00 Matches: 22
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 2 Gaps: 0

US-09-757-415A-3 (1-22) x AAT31051 (1-2733)
QY 1 HisSerGlnMetAlaValHisIleuAlaLysSerIleProLeuArgGlnValThr 20
Db 1342 CACAGCCAGATGGCTGTGCACAAAGCTGCCAAGCATCCCTCTGCGCAGACAGGTAACA 1401
QY 21 ValSer 22
Db 1402 GTGTCT 1407

RESULT 11
AAQ10448
ID AAQ10448 standard; cDNA; 2856 BP.
XX AC
XX AAQ10448;
XX DT 25-MAR-2003 (revised)
XX DT 17-APR-1991 (first entry)
XX XX
XX DE Human basic fibroblast growth factor.
XX KW FGF; angiogenesis; neuron growth; ss.
XX OS Homo sapiens.

```

XX Key Location/Qualifiers
KW sig_peptide 529..591
OS /*tag= a
XX mat_peptide 592..2721
FT /*tag= b
FT /*product= "bFGF-R"
XX WO9100916-A.
XX 24-JAN-1991.
XX 06-JUL-1989; 89US-00377003.
XX 06-JUL-1989; 89US-00377003.
XX (REGC) UNIV CALIFORNIA.
XX Williams LT, Johnson DE, Lee PE;
XX WPI; 1991-051340/07.
XX P-PSDB; AAR10649.
XX Fibroblast growth factor receptors - and clones encoding them, sol. forms
XX of receptors and antibodies to them useful for diagnostic studies and
XX therapy.
XX Disclosure; Fig 4; 84pp; English.
XX The sequence was obtd. from a clone isolated from a huma ubilical vein
XX endothelial cDNA library and a human placenta cDNA library using a probe
XX based on the hfg cDNA sequence. Four types of positive clones were
XX isolated, divided into two classes based on their restriction patterns:
XX h2 and h3 (membrane spanning recep- tors), and h4 and h5 (secreted
XX forms). The sequence here is that of one of the larger clones (h2/h3).
XX The sequence of bases 2857- 3612 appears to be absent from the
XX specification, with the 3' terminal of bases 3613-3636 given as:
XX gctgggagcaggtcgcaaggga. The DNA encodes a basic fibroblast growth factor
XX receptor which can be used for tissue repair or development and for
XX neuron maintenance. The DNA and protein and antibodies prep'd. from it
XX can be used to treat diseases resulting from aberrant FGF expression and
XX for elucidating the mechanism of FGF and identify- ing agonists for the
XX prevention of angiogenesis and control of tumours. See also AAQ10448.
XX (Updated on 25-MAR-2003 to correct PA field.)
XX SQ Sequence 2856 BP; 631 A; 872 C; 836 G; 517 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 9,08e-09 Length: 2856
Score: 107.00 Matches: 22
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 2 Gaps: 0
US-09-757-415A-3 (1-22) x AAQ10448 (1-2856)
QY 1 HisSerGlnMetAlaValHisLysLeuAlaLysSerIleProLeuArgArgGlnValThr 20
Db 1480 CACAGCCAGATGGCTGTGCACAAAGCTGGCCAGAGCATCCCTCTCCGACAGGTACACA 1539
QY 21 Valser 22
Db 1540 GTGTCT 1545
RESULT 12
AAQ20914
ID AAQ20914 standard; cDNA; 3328 BP.
XX
AC AAQ20914;
XX 03-JUL-1992 (first entry)
DT
XX Human bFGF receptor sequence.

XX Basic fibroblast growth factor; fms like gene; ss.
XX Homo sapiens.
XX Key Location/Qualifiers
FT CDS 229..2425
FT /*tag= a
FT /*product
FT /*note= "h bFGF receptor"
XX JP04036184-A.
XX 06-FEB-1992.
XX 01-JUN-1990; 90JP-00141490.
XX 01-JUN-1990; 90JP-00141490.
XX (MITK) MITSUI TOATSU CHEM INC.
XX WPI; 1992-092893/12.
XX P-PSDB; AAR21686.
XX Human b-FGF receptor gene - contains recombinant DNA, chromosome DNA or
XX cDNA base sequence encoding polypeptide.
XX Claim 1; Fig 1; 10pp; Japanese.
XX The human basic fibroblast growth factor (b-FGF) receptor sequence was
XX obtd. from a human placenta cDNA library in lambda gt11. The library was
XX screened with two probes (AAQ24634,5) which were derived from the known
XX FGF sequence (fms like gene). Positive clones were transformed into E.
XX coli Y1090 strain, and recombinant phage were sequenced
XX SQ Sequence 3328 BP; 777 A; 946 C; 902 G; 703 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 1.1e-08 Length: 3328
Score: 107.00 Matches: 22
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 2 Gaps: 0
US-09-757-415A-3 (1-22) x AAQ20914 (1-3328)
QY 1 HisSerGlnMetAlaValHisLysLeuAlaLysSerIleProLeuArgArgGlnValThr 20
Db 1180 CACAGCCAGATGGCTGTGCACAAAGCTGGCCAGAGCATCCCTCTCCGACAGGTACACA 1239
QY 21 Valser 22
Db 1240 GTGTCT 1245
RESULT 13
ABZ35187
ID ABZ35187 standard; cDNA; 3328 BP.
XX
AC ABZ35187;
XX 05-FEB-2003 (first entry)
DT
XX Human gene expression profile polynucleotide SEQ ID NO 299.
XX Human; artery; endothelium; umbilical; vein; aorta; pulmonary artery;
XX bronchial epithelium; prostate; muscle; lung fibroblast; osteoblast;
XX tumour; microarray; genome mapping; antibiotic; antiviral; antifungal;
XX Gene expression; Gene; ss.
XX Homo sapiens.
XX WO200274979-A2.
PN

```
XX PD 26-SEP-2002.
XX KW ss; cytostatic; farnesyl transferase inhibitor; gene expression;
XX PF quinolone; leukemia; cancer.
XX OS Homo sapiens.
XX EN WO2003038129-A2.
XX XX
XX PD 08-MAY-2003.
XX XX
XX PF 30-OCT-2002; 2002WO-US034784.
XX XX
XX PR 30-OCT-2001; 2001US-0338997P.
XX PR 30-OCT-2001; 2001US-0340881P.
XX PR 30-OCT-2001; 2001US-0340938P.
XX PR 30-OCT-2001; 2001US-0341012P.
XX XX
XX PA (ORTH ) ORTHO CLINICAL DIAGNOSTICS INC.
XX XX
XX PI Raponi M;
XX XX
XX XX WPI; 2003-513497/48.
XX XX
XX PT Determining whether a patient will respond to treatment with a farnesyl
XX PT transferase inhibitor, by analyzing the expression of gene that is
XX PT differentially modulated in the presence of the inhibitor.
XX XX
XX PS Disclosure; SEQ ID NO 98; 346pp; English.
XX XX
XX CC The invention relates to a method of determining whether a patient will
XX CC respond to treatment with a farnesyl transferase inhibitor (FTI), by
XX CC analyzing the expression of gene that is differentially modulated in the
XX CC presence of an FTI. The method is useful for determining whether a
XX CC patient will respond to treatment with a FTI such as (2)-6-[amino(4-
XX CC chlorophenyl)(1-methyl-1H-imidazol-5-yl)methyl]-4-(3-chlorophenyl)-1-
XX CC methyl-2-(1H)quinoline, monitoring the therapy of a patient, treating a
XX CC patient with leukemia with FTI if the analysis indicates that the patient
XX CC will respond. This sequence corresponds to a gene whose expression may be
XX CC modulated in the presence of FTI.
XX XX
XX SQ Sequence 3328 BP; 777 A; 946 C; 902 G; 703 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1.1e-08 Length: 3328
Score: 107.00 Matches: 22
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-09-757-415A-3 (1-22) x ABZ35187 (1-3328)
QY 1 HisSerGlnMetAlaValHisLysLeuAlaLysSerIleProLeuArgGlnValThr 20
Db 1180 CACAGCCAGATGGCTGTGCACAGCTGCCAAGAGCATCCCTCTCCGACAGGTAACA 1239

QY 21 ValSer 22
Db 1240 GTGTCT 1245

RESULT 14
ADE84879
ID ADE84879 standard; DNA; 3328 BP.
XX
XX AC ADE84879;
XX XX
XX DT 29-JAN-2004 (first entry)
XX XX
XX DE Farnesyl transferase inhibitor modulated leukemia associated gene #98.

XX XX
XX PD 26-SEP-2002.
XX KW ss; cytostatic; farnesyl transferase inhibitor; gene expression;
XX PF quinolone; leukemia; cancer.
XX OS Homo sapiens.
XX EN WO2003038129-A2.
XX XX
XX PD 08-MAY-2003.
XX XX
XX PF 30-OCT-2002; 2002WO-US034784.
XX XX
XX PR 30-OCT-2001; 2001US-0338997P.
XX PR 30-OCT-2001; 2001US-0340881P.
XX PR 30-OCT-2001; 2001US-0340938P.
XX PR 30-OCT-2001; 2001US-0341012P.
XX XX
XX PA (ORTH ) ORTHO CLINICAL DIAGNOSTICS INC.
XX XX
XX PI Wan J, Wang Y;
XX XX
XX PF WPI; 2002-740862/80.
XX XX
XX PT New gene expression profile generated from primary, endothelial,
XX PT epithelial, and muscle cell types, useful for identifying disease
XX PT pathologies involving alterations of gene expression, e.g. cancer.
XX XX
XX PS Claim 10; Page 454-455; 850pp; English.
XX XX
XX CC The invention relates to a gene expression profile comprising one or more
XX CC genes (ABZ34869-ABZ35692) and generated from a cell type. The cell type
XX CC is a coronary artery endothelium, umbilical artery or vein endothelium,
XX CC aortic endothelium, dermal microvascular endothelium, pulmonary artery
XX CC endothelium, myometrium microvascular endothelium, keratinocyte
XX CC epithelium, bronchial epithelium, mammary epithelium, prostate
XX CC epithelium, renal cortical epithelium, renal proximal tubule epithelium,
XX CC small airway epithelium, renal epithelium, umbilical artery smooth
XX CC muscle, neonatal dermal fibroblast, pulmonary artery smooth muscle,
XX CC dermal fibroblast, neural progenitor cells, skeletal muscle, astrocytes,
XX CC aortic smooth muscle, mesangial cells, coronary artery smooth muscle,
XX CC bronchial smooth muscle, uterine smooth muscle, lung fibroblast,
XX CC osteoblasts or prostate stromal cell. The gene expression profile is used
XX CC for determining the level of RNA expression for a sample, determining the
XX CC phenotype of a cell and distinguishing cell types. The gene or a protein
XX CC expression profile is useful in identifying disease pathologies involving
XX CC alterations of gene expression. The assessment of expression profiles may
XX CC provide meaningful information with respect to tumour type and stage,
XX CC treatment methods, and prognosis. The gene or protein expression profile
XX CC may also be used for creating microarrays. The microarray is useful for
XX CC genetic and physical mapping of genomes, DNA sequencing, genetic or
XX CC medical diagnosis, genotyping of organisms, confirming cell or tissue
XX CC identifications and in identifying promising antibiotics, antiviral or
XX CC antifungal agents
XX XX
XX SQ Sequence 3328 BP; 777 A; 946 C; 902 G; 703 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1.1e-08 Length: 3328
Score: 107.00 Matches: 22
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-09-757-415A-3 (1-22) x ABZ35187 (1-3328)
QY 1 HisSerGlnMetAlaValHisLysLeuAlaLysSerIleProLeuArgGlnValThr 20
Db 1180 CACAGCCAGATGGCTGTGCACAGCTGCCAAGAGCATCCCTCTCCGACAGGTAACA 1239

QY 21 ValSer 22
Db 1240 GTGTCT 1245

RESULT 14
ADE84879
ID ADE84879 standard; DNA; 3328 BP.
XX
XX AC ADE84879;
XX XX
XX DT 29-JAN-2004 (first entry)
XX XX
XX DE Farnesyl transferase inhibitor modulated leukemia associated gene #98.
```

```
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 10..12472
FT /tag= a
FT misc_feature 1287..1288
FT /tag= b
FT /note= "t(8;13) breakpoint"
XX
XX WO9935159-A1.
XX
XX 15-JUL-1999.
XX
XX 07-JAN-1999; 99WO-US000365.
XX
XX 08-JAN-1998; 98US-00004688.
XX
XX (BGMH ) BRIGHAM & WOMENS HOSPITAL.
XX
XX Fletcher JA, Xiao S;
XX
XX WPI; 1999-430381/36.
XX P-PSDB; AAY06458.
XX
XX Oncogene, designated ZNF198-FGFR1, useful in diagnosis of stem cell
XX leukemia and lymphoma syndrome.
XX
XX Disclosure; Page 69-70; 76pp; English.
XX
XX This is the nucleotide sequence of human fibroblast growth factor
XX receptor 1 (FGFR1) cDNA. The FGFR1 gene is located on chromosome 8. A
XX newly identified oncogene, ZNF198-FGFR1 (see AAX87177), incorporates a
XX FGFR1 tyrosine kinase domain fused through t(8;13) translocation to
XX ZNF198, a novel zinc finger gene of chromosome 13, and is associated with
XX stem cell leukaemia/lymphoma syndrome (SCLL). Molecular characterisation
XX of the oncogene and the encoded oncoprotein (see AAY06456) provides the
XX ability to identify patients with t(8;13)(p11;q11-12) syndrome prior to
XX onset of myeloid leukaemia. The oncogene can be identified with primers
XX and probes capable of hybridizing to ZNF198 or FGFR1 tyrosine kinase
XX domain coding sequences. Cells containing ZNF198-FGFR1 can be located
XX using a radiolabeled antibody that recognizes the fusion junction locus.
XX Antibodies that recognize this fusion junction can also be used to
XX deliver toxic substances to patients containing the ZNF198-FGFR1 protein.
XX This is useful for treatment of SCLL. Antisense nucleotides and ribozymes
XX targeted to ZNF198-FGFR1 oncogene are useful for reducing expression or
XX disrupting translation of the oncogene. (All claimed)
XX
XX Sequence 3343 BP; 766 A; 957 C; 911 G; 709 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1.11e-08 Length: 3343
Score: 107.00 Matches: 22
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 2 Gaps: 0

US-09-757-415A-3 (1-22) x AAX87179 (1-3343)

Qy 1 HisSerGlnMetAlaValHisLysLeuAlaLysSerIleProIeuArgGlnValThr 20
Db 1228 CACAGCCAGATGGCTGTGTGCACAGCTGGCCAGAGAGCATCCTCTGCGCAGACAGGTAACA 1287
Qy 21 ValSer 22
Db 1288 GTGTCT 1293

Search completed: August 8, 2004, 21:31:41
Job time : 96.8421 secs
```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 8, 2004, 21:03:30 ; Search time 610.355 Seconds
(without alignments)
1076.370 Million cell updates/sec

Title: US-09-757-415A-3
Perfect score: 107
Sequence: 1 HSQMAVHKLAKSLPRLRQVTVS 22

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-Q/cgn2 1/USPTO spoof/US09757415/runat 02082004 101231 6697/app query.fasta_1.526
-DB=EST -QFMT=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -FRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09757415 @CGN 1.1 4385 @runat 02082004 101231 6697 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

- Database :
- 1: em_estba:**
 - 2: em_esthum:**
 - 3: em_estin:**
 - 4: em_estmu:**
 - 5: em_estov:**
 - 6: em_estpl:**
 - 7: em_estro:**
 - 8: em_hc:**
 - 9: gb_est1:**
 - 10: gb_est2:**
 - 11: gb_hc:**
 - 12: gb_est3:**
 - 13: gb_est4:**
 - 14: gb_est5:**
 - 15: em_estfun:**
 - 16: em_estom:**
 - 17: em_gss_hum:**
 - 18: em_gss_inv:**
 - 19: em_gss_pln:**
 - 20: em_gss_vrt:**
 - 21: em_gss_fun:**
 - 22: em_gss_mam:**
 - 23: em_gss_mus:**
 - 24: em_gss_pro:**
 - 25: em_gss_rod:**
 - 26: em_gss_phg:**
 - 27: em_gss_vrl:**
 - 28: gb_gss1:**

29: gb_gss2:**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description	
1	107	100.0	257	10	AW446321	AW446321 84573 MAR	
2	107	100.0	305	13	BY363360	BY363360 BY363360	
3	107	100.0	327	13	BY362999	BY362999 BY362999	
4	107	100.0	343	10	AW446323	AW446323 84979 MAR	
5	107	100.0	377	14	M78197	M78197 EST01790 Su	
6	107	100.0	397	14	CB772368	CB772368 AMGNNUC:T	
7	107	100.0	399	14	CB706389	CB706389 AMGNNUC:T	
8	107	100.0	415	14	CB800339	CB800339 AMGNNUC:T	
9	107	100.0	419	10	BF922260	BF922260 QV2-NT014	
c	10	107	100.0	451	14	CB787182	CB787182 AMGNNUC:N
11	107	100.0	474	12	EM088352	EM088352 501802 MA	
12	107	100.0	475	10	AW660624	AW660624 99589 MAR	
13	107	100.0	483	14	CD550008	CD550008 B0305F05-	
14	107	100.0	498	14	CA395172	CA395172 CS61B08.Y	
15	107	100.0	514	10	BE531255	BE531255 601231076	
16	107	100.0	519	9	AA637508	AA637508 vr32d08.r	
17	107	100.0	529	9	AU127333	AU127333 AU127333	
18	107	100.0	545	9	AA388928	AA388928 mp15g07.r	
19	107	100.0	546	9	AU126897	AU126897 AU126897	
20	107	100.0	570	9	AA067236	AA067236 mm30d09.r	
c	21	107	100.0	580	10	BF888342	BF888342 MR4-TN010
22	107	100.0	598	14	CF125564	CF125564 UI-HF-EL0	
23	107	100.0	600	12	BG061112	BG061112 2014-36 M	
24	107	100.0	605	14	CA848058	CA848058 iq42b06.Y	
25	107	100.0	616	13	BU054527	BU054527 UI-M-FD0-	
26	107	100.0	623	10	AK320452	AK320452 uc019b01.Y	
27	107	100.0	629	10	BF094753	BF094753 MR3-UT005	
28	107	100.0	632	14	CF750302	CF750302 UI-M-H10-	
29	107	100.0	648	14	CF125642	CF125642 UI-HF-EL0	
30	107	100.0	652	10	BF688338	BF688338 292100 BA	
31	107	100.0	665	12	BF653629	BF653629 603300272	
32	107	100.0	712	10	BF539113	BF539113 602051471	
33	107	100.0	721	12	BF452957	BF452957 603170110	
34	107	100.0	729	14	CA750867	CA750867 UI-M-FD0-	
35	107	100.0	740	9	AU131387	AU131387 AU131387	
36	107	100.0	747	9	AU130342	AU130342 AU130342	
37	107	100.0	752	9	AU134076	AU134076 AU134076	
38	107	100.0	800	13	BUS44807	BUS44807 AGENCOURT	
39	107	100.0	801	14	CD655176	CD655176 AGENCOURT	
40	107	100.0	807	9	AU130700	AU130700 AU130700	
41	107	100.0	831	13	BUS55209	BUS55209 AGENCOURT	
42	107	100.0	834	12	BF753909	BF753909 603022913	
43	107	100.0	867	9	AU119392	AU119392 AU119392	
44	107	100.0	874	10	BF160004	BF160004 601768184	
45	107	100.0	874	12	BF688445	BF688445 603316382	

ALIGNMENTS

RESULT 1
AW446321
LOCUS 84973 MARC 1BOV Bos taurus cDNA 5', mRNA linear EST 25-APR-2001
DEFINITION 84973 MARC 1BOV Bos taurus cDNA 5', mRNA sequence.
ACCESSION AW446321
VERSION AW446321.1 GI:6988107
KEYWORDS EST.
SOURCE Bos taurus (cow)
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
REFERENCE 1 (bases 1 to 257)

AUTHORS Smith, T. P. L., Grosse, W. M., Freking, B. A., Roberts, A. J., Stone, R. T., Casas, E., Wray, J. E., White, J., Cho, J., Fahrenkrug, S. C., Bennett, G. L., Heaton, M. P., Laegreid, W. W., Rohrer, G. A., Chitko-Mckown, C. G., Perteau, G., Holt, I., Karamycheva, S., Liang, F., Quackenbush, J. G., and Keefe, J. W.

TITLE Sequence evaluation of four pooled-tissue normalized bovine cDNA libraries and construction of a gene index for cattle

JOURNAL Genome Res. 11 (4), 626-630 (2001)

MEDLINE 21180013

PUBMED 11282978

COMMENT Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smithemail.marc.usda.gov

Single pass sequencing. Bases called and trimmed with phred v0.980904.e. Vector identified by cross_match with the -minscore 20 and -mismatch 12 options.

PCR Primers
FORWARD: AGGAACAGCTATGACCAT
BACKWARD: GTTTCACGTCAGGAGG
Plate: 53 row: A column: 6
Seq primer: ATTAGGTGACACTATAG.

FEATURES Location/Qualifiers
1..257
/organism="Bos taurus"
/mol_type="mRNA"
/db_xref="taxon:9913"
/tissue_type="pooled"
/lab_host="DH10B"
/clone_lib="MARC 1BOV"
note="Vector: pCMV SPORT6; Site 1: NotI; Site 2: SalI;
Library made from pooled tissue from lymph node, ovary,
fat, hypothalamus, and pituitary."

ORIGIN

Alignment Scores:	Length:	257
Pred. No.:	3.16e-07	
Score:	107.00	22
Percent Similarity:	100.00%	Conservative: 0
Best Local Similarity:	100.00%	Mismatches: 0
Query Match:	100.00%	Indels: 0
DB:	10	Gaps: 0

US-09-757-415A-3 (1-22) x AW446321 (1-257)

Qy 1 HisSerGlnMetAlaValHisLysLeuAlaLysSerIleProLeuArgArgGlnValThr 20
|||||
Db 47 CACAGCCAGATGGCGTGCACAGCTGGCCAGAGAGCATCCCTCTGCCACAGAGTAA 106

Qy 21 ValSer 22
|||||
Db 107 GTGTGC 112

RESULT 2
BY363360 305 bp mRNA linear EST 12-DEC-2002
LOCUS BY363360
DEFINITION musculus cDNA clone C230084A06 3', mRNA sequence.
ACCESSION BY363360
VERSION BY363360.1 GI:26592848
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 305)
Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S.,
Nikaido, I., Osato, N., Saito, R., Suzuki, H., Yamanaka, I.,
Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A.,
Schonbach, C., Gojohori, T., Baldarelli, R., Hill, D. P., Sult, C.,
Hume, D. A., Quackenbush, J., Schriml, L. M., Kanapin, A., Matsuda, H.,

TITLE

Analysis of the mouse transcriptome based on functional annotation

of 60,770 full-length cDNAs

Nature 420, 563-573 (2002)

22354683

12466851

Contact: Yoshihide Hayashizaki

Laboratory for Genome Exploration Research Group, RIKEN Genomic

Sciences Center (GSC), Yokohama Institute

The Institute of Physical and Chemical Research (RIKEN)

1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

Tel: 81-45-503-9222

Fax: 81-45-503-9216

Email: genome-res@gsc.riken.go.jp

URL: http://genome.gsc.riken.go.jp/

Alizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S.,

Hirozane, T., Imotani, K., Ishii, Y., Itoh, M., Kawai, J., Konno, H.,

Miyazaki, A., Murata, M., Nakamura, M., Nomura, K., Numazaki, R.,

Ohno, M., Sakai, K., Sakazume, N., Sasaki, D., Sato, K., Shibata, K.,

Shiraki, T., Tagami, M., Waki, K., Watahiki, A., Muramatsu, M. and

Hayashizaki, Y. Direct Submission

Computational Analysis of Full-length Mouse cDNAs Compared with

Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)

Normalization and subtraction of cap-trapper-selected cDNAs to

prepare full-length cDNA libraries for rapid discovery of new

genes. Genome Res. 10 (10), 1617-1630 (2000)

RIKEN integrated sequence analysis (RISA) system--384-format

sequencing pipeline with 384 multicapillary sequencer. Genome Res.

10 (11), 1757-1771 (2000)

Computer-based methods for the mouse full-length cDNA

encyclopedia: real-time sequence clustering for construction of a

nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)

cDNA library was prepared and sequenced in Mouse Genome

Encyclopedia Project of Genome Exploration Research Group in Riken

Genomic Sciences Center and Genome Science Laboratory in Riken.

Division of Experimental Animal Research in Riken contributed to

prepare mouse tissues.

Please visit our web site (<http://genome.gsc.riken.go.jp>) for

further details.

Location/Qualifiers

1..305

/organism="Mus musculus"

/mol_type="mRNA"

/db_xref="taxon:10090"

/clone="C230084A06"

/tissue_type="cerebellum"

/dev_stage="0 day neonate"

/lab_host="DH10B"

/clone_lib="RIKEN full-length enriched, 0 day neonate

cerebellum"

note="Site 1: SalI; Site 2: BamHI; cDNA library was

prepared and sequenced in Mouse Genome Encyclopedia

FEATURES

source

Score: 107.00 Matches: 22
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 13 Gaps: 0

US-09-757-415A-3 (1-22) x BY362999 (1-327)

QY 1 HisSerGlnMetAlaValHisLysLeuAlaLysSerIleProLeuArgGlnValThr 20
 DB 168 CATAGCCAGATGGCTGTGCACAAAGCTGCCAAGAGCATCCCTCTGCCGACAGGTAACA 227

QY 21 ValSer 22

DB 228 GTGTCA 233

RESULT 4

AW446323

LOCUS 343 bp mRNA linear EST 25-APR-2001

DEFINITION 84979 MARC 1BOV Bos taurus cDNA 5', mRNA sequence.

ACCESSION AW446323

VERSION AW446323.1 GI:6988109

KEYWORDS EST.

SOURCE Bos taurus (cow)

ORGANISM

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 Bovidae; Bovinae; Bos.

AUTHORS Smith,T.P.L., Grosse,W.M., Freking,B.A., Roberts,A.J., Stone,R.T.,
 Casas,E., Wray,J.E., White,J., Cho,J., Fahrenkrug,S.C.,
 Bennett,G.L., Heaton,M.P., Laegreid,W.W., Rohrer,G.A.,
 Chitko-McKown,C.G., Pertea,G., Holt,I., Karamycheva,S., Liang,F.,
 Quackenbush,J. and Keele,J.W.

TITLE Sequence evaluation of four pooled-tissue normalized bovine cDNA
 libraries and construction of a gene index for cattle
 Genome Res. 11 (4), 626-630 (2001)

JOURNAL

MEDLINE

PUBMED

COMMENT

Contact: Smith TPL
 USDA, ARS, US Meat Animal Research Center
 PO Box 166, Clay Center, NE 68933-0166, USA
 Tel: 402 762 4366
 Fax: 402 762 4390

Email: smith@email.marc.usda.gov

Single pass sequencing. Bases called and trimmed with phred
 v0.980904.e. Vector identified by cross_match with the -minscore 20
 and -minmatch 12 options.

PCR Primers

FORWARD: AGGAACAGCTATGACCAT

BACKWARD: GTTTCCTCAGTCAGCAGC

Plate: 53 row: B column: 6

Seq primer: ATTAGGTGACACTATAG.

Location/Qualifiers

FEATURES

source

1..343

/organism="Bos taurus"

/mol_type="mRNA"

/db_xref="taxon:9913"

/tissue_type="pooled"

/lab_host="DH10B"

/clone_lib="WARC 1BOV"

/note="vector; pCMV SPOR6; Site 1: NotI; Site 2: SalI;

Library made from pooled tissue from lymph node, ovary,

fat, hypothalamus, and pituitary."

ORIGIN

Alignment Scores:

Pred. No.: 4.61e-07 Length: 343
 Score: 107.00 Matches: 22
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 10 Gaps: 0

US-09-757-415A-3 (1-22) x AW446323 (1-343)

QY 1 HisSerGlnMetAlaValHisLysLeuAlaLysSerIleProLeuArgGlnValThr 20
 DB 47 CACAGCCAGATGGCGGTGCACAAAGCTGCCAAGAGCATCCCTCTGCCGACAGGTAACA 106

QY 21 ValSer 22

DB 107 GTGTGC 112

RESULT 5

M78197

LOCUS

DEFINITION

M78197 377 bp mRNA linear EST 26-MAY-1992
 EST01790 Subtracted Hippocampus, Stratagene (cat. #936205) Homo
 sapiens cDNA clone HPCPQ03 similar to Fibroblast growth factor
 receptor, mRNA sequence.

ACCESSION M78197

VERSION M78197.1 GI:273934

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS Adams,M.D., Dubnick,M., Kerlavage,A.R., Moreno,R., Kelley,J.M.,
 Utterback,T.R., Nagle,J.W., Fields,C. and Venter,J.C.

TITLE Sequence identification of 2,375 human brain genes

JOURNAL Nature 355, 632-634 (1992)

MEDLINE 92188112

PUBMED 1538749

COMMENT

Contact: Kerlavage, AR

Bioinformatics

The Institute for Genomic Research

9712 Medical Center Drive, Rockville, MD 20850 USA

Tel: 3018699056

Fax: 3018699423

Email: arkerlav@tigr.org

Seq primer: M13 Forward.

Location/Qualifiers

FEATURES

source

1..377

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="ATCC (inhost):78796"

/db_xref="GDB:D0S1193E"

/db_xref="taxon:9606"

/clone="HPCPQ03"

/clone_lib="Subtracted Hippocampus, Stratagene (cat.
 #936205)"

/note="Vector: lambdaZAP-II; The hippocampus library (#4
 above) was subtracted with a fibroblast cell line cDNA
 library (Stratagene cat. #936209; WI38 lung fibroblast
 cell line; oligo-dT + random primed cDNA synthesis;
 lambdaZAP-II vector, 1.0kb average insert size.) by the
 method of Sive & St. John (Nucl. Acids Res. 16:10937,
 1988)."

ORIGIN

Alignment Scores:

Pred. No.: 5.22e-07 Length: 377
 Score: 107.00 Matches: 22
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 14 Gaps: 0

US-09-757-415A-3 (1-22) x M78197 (1-377)

QY 1 HisSerGlnMetAlaValHisLysLeuAlaLysSerIleProLeuArgGlnValThr 20
 DB 285 CACAGCCAGATGGCTGTGCACAAAGCTGCCAAGAGCATCCCTCTGCCGACAGGTAACA 344

QY 21 ValSer 22

```

Angen, Inc
One Angen Center Drive, Thousand Oaks, CA 91320-1799, USA
Tel: 805 447-4881
Plate: 00001 row: c column: 9.
Location/Qualifiers
1. .399
/organism="Rattus norvegicus"
/mol_type="mRNA"
/db_xref="taxon:10116"
/clone="trcx1-00001-c9"
/tissue_type="choroid plexus brain"
/clone_lib="trcx1 (10261)"
/notes="Vector: pcAV; Site_1: SalI; Site_2: NotI; choroid
plexus brain region"

ORIGIN
Alignment Scores:
Pred. No.: 5.63e-07 Length: 399
Score: 107.00 Matches: 22
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 14 Gaps: 0

US-09-757-415A-3 (1-22) x CB706389 (1-399)

Qy 1 HisSerGlnMetAlaValHisLysLeuAlaLysSerIleProLeuArgArgGlnValThr 20
Db 266 CATAGCAGATGGCTGTGCATAAGCTGGCTAGAGCATCCCTCTCCGCACACAGGTACAA 325

.Qy 21 ValSer 22
|
|
|
Db 326 GTGTCA 331

RESULT 8
CB800339
LOCUS
DEFINITION
CB800339
ACCESSION
CB800339.1 GI:29912517
KEYWORDS
EST.
SOURCE
Rattus norvegicus (Norway rat)
ORGANISM
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE
1 (bases 1 to 415)
AUTHORS
Angen EST Program.
TITLE
Angen Rat EST Program
JOURNAL
Unpublished (2003)
COMMENT
Contact: Dan Fitzpatrick
Angen, Inc
One Angen Center Drive, Thousand Oaks, CA 91320-1799, USA
Tel: 805 447-4881
Plate: 00003 row: f column: 6.
Location/Qualifiers
1. .415
/organism="Rattus norvegicus"
/mol_type="mRNA"
/db_xref="taxon:10116"
/clone="trcp2-00003-f6"
/tissue_type="choroid plexus brain"
/clone_lib="trcp2 (10289)"
/notes="Vector: C6KGF7L; Site_1: SalI; Site_2: NotI;
choroid plexus brain region"

ORIGIN
Alignment Scores:
Pred. No.: 5.93e-07 Length: 415
Score: 107.00 Matches: 22
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0

```

Query Match: 100.00% Indels: 0
DB: 14 Gaps: 0

US-09-757-415A-3 (1-22) x CB800339 (1-415)

QY 1 HisserGlnMetAlaValHislysleuAlaValSerIleProLeuArgGlnValThr 20
|||||
DB 279 CATAGCCAGATGGCTGTCATTAAGCTGGCTAAGAGCATCCCTCTCCGACAGAGTAACA 338
|||||
QY 21 ValSer 22
|||||
DB 339 GTGTCA 344

RESULT 9
BF922260/c
LOCUS
DEFINITION QV2-NT0147-191100-483-a06 NT0147 Homo sapiens cDNA, mRNA sequence.
ACCESSION BF922260
VERSION BF922260.1 GI:12318148
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 419)
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W.Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
PUBMED 10737800
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=QV2&st2=QV2-NT0147-191100-483-a06&st3=2000-11-19&st4=1)
Seq primer: puc 18 forward
High quality sequence stop: 419.
Location/Qualifiers
1..419
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="NT0147"
/note="Organ: nervous tumor; Vector: puc18; Site: 1: Smat; Site 2: Smat; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

FEATURES
source
1..419
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="NT0147"
/note="Organ: nervous tumor; Vector: puc18; Site: 1: Smat; Site 2: Smat; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

ORIGIN
Alignment Scores:
Pred. No.: 6e-07 Length: 419
Score: 107.00 Matches: 22
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 10 Gaps: 0

US-09-757-415A-3 (1-22) x BF922260 (1-419)

QY 1 HisserGlnMetAlaValHislysleuAlaValSerIleProLeuArgGlnValThr 20
|||||
DB 183 CACAGCCAGATGGCTGTCACCAAGCTGCCAAGAGCATCCCTCTCCGACAGAGTAACA 124
|||||
QY 21 ValSer 22
|||||
DB 123 GTGTCT 118

RESULT 10
CB787182
LOCUS
DEFINITION AWGNNUC:NRHY4-00054-B8-A W Rat hypothalamus (10464) Rattus norvegicus cDNA clone nrhy4-00054-B8 5', mRNA sequence.
ACCESSION CB787182
VERSION CB787182.1 GI:29875573
KEYWORDS EST.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
REFERENCE 1 (bases 1 to 451)
AUTHORS Amgen EST Program.
TITLES Amgen Rat EST Program
JOURNAL Unpublished (2003)
COMMENT Contact: Dan Fitzpatrick
Amgen, Inc
One Amgen Center Drive, Thousand Oaks, CA 91320-1799, USA
Tel: 805 447-4881
Plate: 00054 row: b column: 8.
Location/Qualifiers
1..451
/organism="Rattus norvegicus"
/mol_type="mRNA"
/db_xref="taxon:10116"
/clone="nrhy4-00054-B8"
/clone_lib="W Rat hypothalamus (10464)"
/note="Vector: pSPORT1; Site 1: SalI; Site 2: NotI; W Rat hypothalamus adult female Wistar rat avg. Insert size 2.3 kb fraction 6 and 7"

FEATURES
source
1..451
/organism="Rattus norvegicus"
/mol_type="mRNA"
/db_xref="taxon:10116"
/clone="nrhy4-00054-B8"
/clone_lib="W Rat hypothalamus (10464)"
/note="Vector: pSPORT1; Site 1: SalI; Site 2: NotI; W Rat hypothalamus adult female Wistar rat avg. Insert size 2.3 kb fraction 6 and 7"

ORIGIN
Alignment Scores:
Pred. No.: 6.61e-07 Length: 451
Score: 107.00 Matches: 22
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 14 Gaps: 0

US-09-757-415A-3 (1-22) x CB787182 (1-451)

QY 1 HisserGlnMetAlaValHislysleuAlaValSerIleProLeuArgGlnValThr 20
|||||
DB 204 CATAGCCAGATGGCTGTCATTAAGCTGGCTAAGAGCATCCCTCTCCGACAGAGTAACA 263
|||||
QY 21 ValSer 22
|||||
DB 264 GTGTCA 269

RESULT 11
BM088352
LOCUS
DEFINITION 501802 MARC 2B0V Bos taurus cDNA 5', mRNA sequence.
ACCESSION BM088352
VERSION BM088352.1 GI:16998980
KEYWORDS EST.
SOURCE Bos taurus (cow)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
1 (Bases 1 to 474)
REFERENCE
AUTHORS
Smith, I.P.L., Grosse, W.M., Freking, B.A., Roberts, A.J., Stone, R.T.,
Casas, E., Wray, J.E., White, J., Cho, J., Fahrenkrug, S.C.,
Bennett, G.L., Heaton, M.P., Laegreid, W.M., Rohrer, G.A.,
Chitko-McKown, C.G., Pettea, G., Holt, I., Karamycheva, S., Liang, F.,
Quackenbush, J. and Keefe, J.W.
Sequence evaluation of four pooled-tissue normalized bovine cDNA
libraries and construction of a gene index for cattle
Genome Res. 11 (4), 626-630 (2001)
PUBMED
21180013
COMMENT
Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called and alt trimmed with phred
v0.980904.e. Vector identified by cross_match with the -minscore 18
and -minmatch 12 options.
PCR Primers
FORWARD: AGGAACAGCTATGACCAT
BACKWARD: GTTTCACGTCACGACG
Plate: 140 row: A column: 21
Seq primer: ATTAGGTGACACTATAG.
Location/Qualifiers
FEATURES
source
1..474
/organism="Bos taurus"
/mol_type="mRNA"
/db_xref="taxon:9913"
/tissue_type="pooled"
/lab_host="DH10B"
/clone_lib="MARC 2BOV"
/note="Vector: pCMV SPORT6; Site 1: NotI; Site 2: SalI;
Library made from pooled tissue from testis, thymus,
semitepidonous muscle, longissimus muscle, pancreas,
adrenal, and endometrium."
ORIGIN
Alignment Scores:
Pred. No.: 7,06e-07 Length: 474
Score: 107.00 Matches: 22
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 12 Gaps: 0
US-09-757-415A-3 (1-22) x BM088352 (1-474)
QY 1 HisSerGlnMetAlaValHisLysLeuAlaLysSerIleProLeuArgGlnValThr 20
|||||
Db 300 CACAGCCAGATGGCGGTGCACAAAGTGGCCCAAGAGCATCCCTCTGCGCAGACAGTAACA 359
QY 21 ValSer 22
Db 360 GTGTCC 365
RESULT 12
LOCUS AW660624 475 bp mRNA linear EST 25-APR-2001
DEFINITION 95589 MARC 1BOV Bos taurus cDNA 5', mRNA sequence.
ACCESSION AW660624
VERSION AW660624.1 GI:7426451
KEYWORDS EST.
SOURCE Bos taurus (cow)
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
1 (Bases 1 to 475)
REFERENCE
AUTHORS Smith, T.P.L., Grosse, W.M., Freking, B.A., Roberts, A.J., Stone, R.T.,

```

```

Casas, E., Wray, J.E., White, J., Cho, J., Fahrenkrug, S.C.,
Bennett, G.L., Heaton, M.P., Laegreid, W.M., Rohrer, G.A.,
Chitko-McKown, C.G., Pettea, G., Holt, I., Karamycheva, S., Liang, F.,
Quackenbush, J. and Keefe, J.W.
Sequence evaluation of four pooled-tissue normalized bovine cDNA
libraries and construction of a gene index for cattle
Genome Res. 11 (4), 626-630 (2001)
PUBMED
21180013
COMMENT
Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called and alt trimmed with phred
v0.980904.e. Vector identified by cross_match with the -minscore 18
and -minmatch 12 options.
PCR Primers
FORWARD: AGGAACAGCTATGACCAT
BACKWARD: GTTTCACGTCACGACG
Plate: 90 row: E column: 20
Seq primer: ATTAGGTGACACTATAG.
Location/Qualifiers
FEATURES
source
1..475
/organism="Bos taurus"
/mol_type="mRNA"
/db_xref="taxon:9913"
/tissue_type="pooled"
/lab_host="DH10B"
/clone_lib="MARC 1BOV"
/note="Vector: pCMV SPORT6; Site 1: NotI; Site 2: SalI;
Library made from pooled tissue from lymph node, ovary,
fat, hypothalamus, and pituitary."
ORIGIN
Alignment Scores:
Pred. No.: 7,08e-07 Length: 475
Score: 107.00 Matches: 22
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 10 Gaps: 0
US-09-757-415A-3 (1-22) x AW660624 (1-475)
QY 1 HisSerGlnMetAlaValHisLysLeuAlaLysSerIleProLeuArgGlnValThr 20
|||||
Db 151 CACAGCCAGATGGCGGTGCACAAAGTGGCCCAAGAGCATCCCTCTGCGCAGACAGTAACA 210
QY 21 ValSer 22
Db 211 GTGTCC 216
RESULT 13
LOCUS CD550008 483 bp mRNA linear EST 11-JUN-2003
DEFINITION B0305F05-5 NIA Mouse E9.5 Whole Embryo cDNA Library (Long) Mus
musculus cDNA clone NIA:B0305F05 IMAGE:30429568 5', mRNA sequence.
ACCESSION CD550008
VERSION CD550008.1 GI:31597750
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (Bases 1 to 483)
REFERENCE
AUTHORS Pao, Y., Ko, N.T., Lim, M.K. and Ko, M.S.H.
T-TITLE Construction of long-transcript enriched cDNA libraries from
submicrogram amounts of total RNAs by a universal PCR amplification
method
Genome Res. 11 (9), 1553-1558 (2001)
PUBMED
21429098
JOURNAL
MEDLINE

```

PUBMED 11544199
 COMMENT Contact: Dawood B. Dudekula
 Laboratory of Genetics
 National Institute on Aging/National Institutes of Health
 333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6920, USA
 Email: cdna@lgsun.grc.nia.nih.gov
 Plate: B0305 row: F column: 05
 Seq primer: M13 Reverse
 High quality sequence scop: 483
 POLYA=No.

FEATURES

source Location/Qualifiers
 1. .483
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="NIA:B0305F05 IMAGE:30429568"
 /tissue type="E9.5 whole embryo"
 /dev stage="whole embryo including extraembryonic tissues at 9.5-days postcoitum"
 /lab_host="DH10B"
 /clone_lib="NIA Mouse E9.5 Whole Embryo cDNA Library (long)"
 /note="Vector: pCMV-SPORT6 (Invitrogen); Site 1: SalI; Site 2: NotI; Mouse cDNA project by the Laboratory of Genetics, National Institute on Aging (NIA), Intramural Research Program, NIH (http://lgsun.grc.nia.nih.gov/cDNA). This is a long-transcript enriched cDNA library (Ref. Genome Res. 11: 1553-1558 (2001). [PMID: 11544199]). Total RNAs were extracted from a pool of 16 embryos at 9.5-days postcoitum. Double-stranded cDNAs were synthesized with an oligo(dT) primer [Invitrogen:
 5'-GCAGTAGTCTAGATCGGAGCGCGCCCTTTTCTTTT-3'] from 6.1 ug of total RNA, treated with T4 DNA polymerase, and purified by ethanol-precipitation. The cDNAs were ligated to lone-linker IL-Sal4, purified by phenol/chloroform, and separated from free linkers by Centricon 100. Then, the cDNAs were amplified by long-range high fidelity PCR using Ex Taq polymerase (Takara) with a primer Sal4-S. The products were purified by phenol/chloroform and Centricon 100. The cDNAs were digested with SalI and NotI enzymes and cloned into SalI/NotI site of pCMV-SPORT6 plasmid vector. The DH10B E. coli host was transformed with the ligation mixture by the standard chemical method. The average insert size is about 3.0kb. The library was constructed by Yulan Piao."

ORIGIN

Alignment Scores:
 Pred. No.: 7.23e-07 Length: 483
 Score: 107.00 Matches: 22
 Percent Similarity: 100.00% Conservatve: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 14 Gaps: 0

US-09-757-415A-3 (1-22) x CD550008 (1-483)

QY 1 HisSerGlnMetAlaValHisLysLeuAlaLysSerIleProLeuArgGlnValThr 20
 Db 269 CATAGCCAGATGGCTGTGCACAACTGGCCAGAGCATCCCTCTGCGACAGAGTAACA 328
 QY 21 ValSer 22
 Db 329 GTGTCA 334

RESULT 14
 CA395172 498 bp mRNA linear EST 06-NOV-2002
 LOCUS cs61b08.y1 Human Retinal pigment epithelium/choroid cDNA
 DEFINITION (Un-normalized, unamplified): cs Homo sapiens cDNA clone cs61b08 5', mRNA sequence.

ACCESSION CA395172
 VERSION CA395172.1 GI:24730401
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 498)

Wistow, G., Bernstein, S.L., Wyatt, M.K., Farris, R.N., Behal, A., Touchman, J.W., Bouffard, G., Smith, D. and Peterson, K.

Expressed sequence tag analysis of human RPE/choroid for the

NEIBank Project: Over 6000 non-redundant transcripts, novel genes

and splice variants

Mol. Vis. 8 (4), 205-220 (2002)

22103460

PubMed

COMMENT

Contact: Wistow G

Section on Molecular Structure and Function

National Eye Institute

6/331 NIH Bethesda, MD 20892-2740, USA

Tel: 301 402 3452

Fax: 301 496 0078

Email: graeme@helix.nih.gov

Plate: 61 row: b column: 08

Seq primer: M13RPI reverse primer (ABI).

Location/Qualifiers

1. .498

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="cs61b08"

/tissue type="RPE/choroid"

/dev stage="Adult"

/lab_host="EMDH10B"

/clone_lib="Human Retinal pigment epithelium/choroid cDNA

(Un-normalized, unamplified): cs"

/note="Organ: Eye; Vector: pCMVSPORT6; Two different donor

eyes (75-80 years old) yielded approximately 600 mg of

dissected RPE/choroid tissue. This in turn yielded 340 ug

of total RNA and 7 ug of mRNA. A directionally cloned cDNA

library in the pCMVSPORT6 vector was constructed at Life

Technologies (Rockville, MD; now part of Invitrogen Corp),

essentially following the protocols of the SuperScript

Plasmid System (Invitrogen Corp).

<http://www.invitrogen.com/>. The library code

designation was cs. For this library, cDNA inserts were

cloned into the NotI/MluI sites of the vector. EST

analysis was performed on the unamplified library at the

NIH Intramural Sequencing Center (NISC)."

ORIGIN

Alignment Scores:
 Pred. No.: 7.53e-07 Length: 498
 Score: 107.00 Matches: 22
 Percent Similarity: 100.00% Conservatve: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 14 Gaps: 0

US-09-757-415A-3 (1-22) x CA395172 (1-498)

QY 1 HisSerGlnMetAlaValHisLysLeuAlaLysSerIleProLeuArgGlnValThr 20
 Db 206 CACAGCCAGATGGCTGTGCACAACTGGCCAGAGCATCCCTCTGCGACAGAGTAACA 265

QY 21 ValSer 22

Db 266 GTGTCT 271

RESULT 15

BE531255

LOCUS

DEFINITION

601231076F1 NCI_CGAP_Mam6 Mus musculus cDNA clone IMAGE:3595050 5',

linear EST 09-AUG-2000

514 bp mRNA

mRNA sequence.
 BE531255
 BE531255.1 GI:9759991
 EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NIH-MGC http://mgi.nci.nih.gov/
 1 (bases 1 to 514)
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 CONTACT: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Jeffrey Green M.D.
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: L1AM8770 row: 9 column: 19
 High quality sequence stop: 513.

FEATURES
 Location/Qualifiers
 1..514
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="FVB/N"
 /db_xref="taxon:10090"
 /clone="IMAGE:3595050"
 /sex="female, virgin"
 /tissue_type="infiltrating ductal carcinoma"
 /dev_stage="5 months"
 /lab_host="DH10B"
 /clone_lib="NCI CGAP Mam6"
 /note="Organ: mammary; Vector: pCMV-SPORT6; Site 1: SalI;
 Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt.
 Library constructed by Life Technologies. Investigator
 providing samples: Jeffrey Green, M.D., NIH"

ORIGIN

Alignment Scores:
 Pred. NO.: 7.85e-07 Length: 514
 Score: 107.00 Matches: 22
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 10 Gaps: 0

US-09-757-415A-3 (1-22) x BE531255 (1-514)

Qy 1 HisSerGlnMetAlaValHisLysLeuAlaLysSerIleProLeuArgGlnValThr 20
 Db 179 CATAGCCAGATGGCTGTGGACAGCTGGCCAGAGCATCCCTCTGCCGACAGAGGTACCA 238
 Qy 21 ValSer 22
 Db 239 GTGTCA 244

Search completed: August 9, 2004, 00:15:16
 Job time : 618.355 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 8, 2004, 21:10:45 ; Search time 16.7895 Seconds

(without alignments)
727.177 Million cell updates/sec

Title: US-09-757-415A-3

Perfect score: 107

Sequence: 1 H5QMAVHLKXSIPLRRQVTS 22

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n_model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09757415/runat_02082004_101232_6719/app_query.fasta_1.526
-DB=Issued_Patents_NA -QFWT=fastap -SUFFIX=p2n.rni -MINMATCH=0.1 -LOOPEL=0
-LOOPEXT=0 -UNITS=Bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFWT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USPR=US09757415 @CGN 1 1 93 @runat_02082004_101232_6719 -NCPU=6 -ICPU=3
-NO MWAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DEEXT=7

Database : Issued Patents NA.*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/2/ina/PCRTUS_COMB.seq.*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	107	100.0	2360	4	US-09-023-655-1288
2	107	100.0	2469	1	US-07-997-333-2
3	107	100.0	2469	1	US-08-459-296-1
4	107	100.0	2469	5	US-07-997-133-2
5	107	100.0	2662	2	US-08-451-822A-14
6	107	100.0	2662	4	US-08-323-430-14
7	107	100.0	2733	1	US-08-371-001-14
8	107	100.0	2733	5	PCT-US96-00331-14
9	107	100.0	3503	1	US-07-631-717A-1
10	107	100.0	3503	1	US-08-166-717D-1
11	83	77.6	1954	1	US-08-471-570-5
12	83	77.6	2676	1	US-08-471-570-7

13	83	77.6	3416	2	US-08-451-822A-15
14	83	77.6	3416	4	US-08-323-430-15
15	81	75.7	2675	1	US-08-070-165F-5
16	81	75.7	2675	2	US-08-885-418-5
17	81	75.7	2681	1	US-08-070-165F-9
18	81	75.7	2681	2	US-08-885-418-9
19	71	66.4	1603	1	US-08-471-570-3
20	71	66.4	2310	1	US-08-471-570-9
21	50	46.7	657	4	US-09-252-991A-1482
22	50	46.7	828	4	US-08-252-991A-1424
23	50	46.7	882	4	US-08-107-532A-2406
24	47	43.9	498	2	US-08-743-637B-168
25	47	43.9	498	3	US-08-526-840B-168
26	47	43.9	611	2	US-08-549-680A-1
27	47	43.9	894	4	US-09-543-681A-2068
28	47	43.9	2823	4	US-09-540-236-725
29	47	43.9	6217	4	US-09-418-817-1
30	47	43.9	7811	2	US-08-549-680A-5
31	47	43.9	48328	4	US-09-596-002-27
32	46	43.0	599	4	US-09-669-751-239
33	46	43.0	1953	4	US-09-107-532A-1867
34	46	43.0	3090	1	US-08-102-942A-3
35	46	43.0	3090	4	US-09-037-179B-3
36	45	42.1	243	4	US-09-489-039A-5655
37	45	42.1	567	4	US-09-489-039A-6174
38	44.5	41.6	1425	4	US-09-489-039A-1509
39	44.5	41.6	2016	4	US-09-489-039A-1492
40	44	41.1	222	4	US-09-863-766A-1
41	44	41.1	1001	4	US-09-641-638-341
42	44	41.1	1001	4	US-09-641-638-342
43	44	41.1	1001	4	US-09-641-638-343
44	44	41.1	1715	3	US-08-757-230A-1
45	44	41.1	1715	4	US-08-700-393-1

ALIGNMENTS

RESULT 1
US-09-023-655-1288
; Sequence 1288, Application US/09023655
; Patent No. 6607879
; GENERAL INFORMATION:
; APPLICANT: Cocks, Benjamin G.
; APPLICANT: Susan G. Stuart
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
; TITLE OF INVENTION: EXPRESSION
; NUMBER OF SEQUENCES: 1508
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/023,655
; FILING DATE: HEREWITH
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0001 US
; TELECOMMUNICATION INFORMATION:

```
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 1288:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2360 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GENBANK
; CLONE: 931386
US-09-023-655-1288

Alignment Scores:
Pred. No.: 1.92e-10 Length: 2360
Score: 107.00 Matches: 22
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-09-757-415A-3 (1-22) x US-09-023-655-1288 (1-2360)
QY 1 HisSerGlnMetAlaValHisLysLeuAlaLysSerIleProLeuArgGlnValThr 20
Db 1114 CACAGCCAGATGGCTGTGCACAAAGCTGCCAAGAGCATCCCTCTGCCGACAGAGTAACA 1173
QY 21 ValSer 22
Db 1174 GTGTCT 1179

RESULT 2
US-07-997-133-2
; Sequence 2, Application US/07997133
; Patent No. 5288855
; GENERAL INFORMATION:
; APPLICANT: Bergonzoni, Laura
; APPLICANT: Mazue, Guy
; APPLICANT: Isacchi, Antonella
; APPLICANT: Roncucci, Romeo
; APPLICANT: Samlientos, Paolo
; TITLE OF INVENTION: Extracellular Form of the Human
; TITLE OF INVENTION: Fibroblast Growth Factor Receptor
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: P.C.
; STREET: 1755 Jefferson Davis Highway, Fourth Floor
; CITY: Arlington
; STATE: Virginia
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/997,133
; FILING DATE: 28-DEC-1992
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/642,755
; FILING DATE: 18-JAN-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Oblon, No. 528885man F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 769-226-0
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)521-4500
; TELEFAX: (703)486-2347
; TELEX: 248855 OPAT-UR
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2469 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; FEATURE:
; NAME/KEY: Coding Sequence
; LOCATION: 1...2466
; OTHER INFORMATION:
US-08-459-296-1

; LENGTH: 2469 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: unknown
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-07-997-133-2

Alignment Scores:
Pred. No.: 2.04e-10 Length: 2469
Score: 107.00 Matches: 22
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 1 Gaps: 0

US-09-757-415A-3 (1-22) x US-07-997-133-2 (1-2469)
QY 1 HisSerGlnMetAlaValHisLysLeuAlaLysSerIleProLeuArgGlnValThr 20
Db 1225 CACAGCCAGATGGCTGTGCACAAAGCTGCCAAGAGCATCCCTCTGCCGACAGAGTAACA 1284
QY 21 ValSer 22
Db 1285 GTGTCT 1290

RESULT 3
US-08-459-296-1
; Sequence 1, Application US/08459296
; Patent No. 5670323
; GENERAL INFORMATION:
; APPLICANT: Gonzalez, Ana-Maria
; APPLICANT: Baird, Andrew
; TITLE OF INVENTION: PROCESS FOR DETECTION OF NEOPLASTIC
; TITLE OF INVENTION: DISEASE
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036/2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/459,296
; FILING DATE: 02-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Mistock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 7573-024
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-790-3090
; TELEFAX: 212-869-8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2469 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; FEATURE:
; NAME/KEY: Coding Sequence
; LOCATION: 1...2466
; OTHER INFORMATION:
US-08-459-296-1
```

Alignment Scores:
Pred. No.: 2,04e-10 Length: 2469
Score: 107.00 Matches: 22
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 1 Gaps: 0

US-09-757-415A-3 (1-22) x US-08-459-296-1 (1-2469)

QY 1 HisSerGlnMetAlaValHisLysLeuAlaLysSerIleProLeuArgGlnValThr 20
Db 1225 CACAGCCAGATGGCTGTGCACAGATGGCCAGAGCATCCCTCTGCCAGACAGGTAAACA 1284

QY 21 ValSer 22
Db 1285 GTGTCT 1290

RESULT 4

US-07-997-133-2
; Sequence 2, Application US/07997133
; GENERAL INFORMATION:
; APPLICANT: Bergonzoni, Laura
; APPLICANT: Mazue, Guy
; APPLICANT: Isacchi, Antonella
; APPLICANT: Roncucci, Romeo
; APPLICANT: Sarmientos, Paolo
; TITLE OF INVENTION: Extracellular Form of the Human
; TITLE OF INVENTION: Fibroblast Growth Factor Receptor
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: P.C.
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; STREET: 1755 Jefferson Davis Highway, Fourth Floor
; CITY: Arlington
; STATE: Virginia
; ZIP: 22202

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/997,133
FILING DATE: 28-DEC-1992
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/642,755
FILING DATE: 18-JAN-1991
ATTORNEY/AGENT INFORMATION:
NAME: Oblon, Norman F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 769-226-0
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)521-4500
TELEFAX: (703)486-2347
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2469 base pairs
TYPE: nucleic acid
STRANDEDNESS: unknown
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)

Alignment Scores:
Pred. No.: 2,04e-10 Length: 2469
Score: 107.00 Matches: 22
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 5 Gaps: 0

US-09-757-415A-3 (1-22) x US-07-997-133-2 (1-2469)

QY 1 HisSerGlnMetAlaValHisLysLeuAlaLysSerIleProLeuArgGlnValThr 20
Db 1225 CACAGCCAGATGGCTGTGCACAGATGGCCAGAGCATCCCTCTGCCAGACAGGTAAACA 1284

QY 21 ValSer 22
Db 1285 GTGTCT 1290

RESULT 5

US-08-451-822A-14
; Sequence 14, Application US/08451822A
; Patent No. 5863888
; GENERAL INFORMATION:
; APPLICANT: Dionne, Craig A
; APPLICANT: Crumley, Greg
; APPLICANT: Jaye, Michael C
; APPLICANT: Schllessinger, Joseph
; TITLE OF INVENTION: Fibroblast Growth Factor Receptors
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rhone-Poulenc Rorer Legal Department
; STREET: 500 Arcola Road
; CITY: Collegeville
; STATE: PA
; COUNTRY: USA
; ZIP: 19426
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/451,822A
FILING DATE: 26-MAY-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/323,430
FILING DATE: 14-OCT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/934,372
FILING DATE: 21-AUG-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/549,587
FILING DATE: 06-JUL-1990
ATTORNEY/AGENT INFORMATION:
NAME: Savitzky, Martin
REGISTRATION NUMBER: 29,699
REFERENCE/DOCKET NUMBER: A0496E
TELECOMMUNICATION INFORMATION:
TELEPHONE: (610) 454-3816
TELEFAX: (610) 454-3808
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 2662 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA

Alignment Scores:
Pred. No.: 2,26e-10 Length: 2662
Score: 107.00 Matches: 22
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 2 Gaps: 0

US-09-757-415A-3 (1-22) x US-08-451-822A-14 (1-2662)

Qy 1 HisSerGlnMetAlaValHisLysLeuAlaLysSerIleProLeuArgGlnValThr 20
Db 1312 CACAGCCAGATGGCTGTGCACCAAGCTGGCCCAAGAGCATCCCTCTGGCGACAGGTAACA 1371

Qy 21 ValSer 22
Db 1372 GTGTCT 1377

RESULT 6

US-08-323-430-14
; Sequence 14, Application US/08323430
; Patent No. 6344546
; GENERAL INFORMATION:
; APPLICANT: Dionne, Craig A
; APPLICANT: Crumley, Greg
; APPLICANT: Jave, Michael C
; APPLICANT: Schlössinger, Joseph
; TITLE OF INVENTION: Fibroblast Growth Factor Receptors
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rhone-Poulenc Rorer Legal Department
; STREET: 500 Arcola Road
; CITY: Collegeville
; STATE: PA
; COUNTRY: USA
; ZIP: 19426
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/323,430
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US to be assigned
; FILING DATE: 21-AUG-1992
; APPLICATION NUMBER: US 07/549,587
; FILING DATE: 06-JUL-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Goodman, Rosanne
; REGISTRATION NUMBER: 32,534
; REFERENCE/DOCKET NUMBER: 10495
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 454-3817
; TELEFAX: (215) 454-3808
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2662 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-323-430-14

Alignment Scores:
Pred. No.: 2,26e-10 Length: 2662
Score: 107.00 Matches: 22
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-09-757-415A-3 (1-22) x US-08-323-430-14 (1-2662)

Qy 1 HisSerGlnMetAlaValHisLysLeuAlaLysSerIleProLeuArgGlnValThr 20
Db 1312 CACAGCCAGATGGCTGTGCACCAAGCTGGCCCAAGAGCATCCCTCTGGCGACAGGTAACA 1371

Qy 21 ValSer 22
Db 1372 GTGTCT 1377

RESULT 7

US-08-371-001-14
; Sequence 14, Application US/08371001
; Patent No. 5783683
; GENERAL INFORMATION:
; APPLICANT: Morrison Ph.D., Richard
; TITLE OF INVENTION: Methods and Composition for Treating
; TITLE OF INVENTION: Tumor Cells
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Gray Cary Ware & Freidenrich
; STREET: 401 "B" Street, Suite 1700
; CITY: San Diego
; STATE: CA
; COUNTRY: USA
; ZIP: 92101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/371,001
; FILING DATE: January 10, 1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Brotman, Harris F.
; REGISTRATION NUMBER: 35,461
; REFERENCE/DOCKET NUMBER: P00095US0
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 699-3630
; TELEFAX: (619) 236-1048
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2733 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 210..467
; OTHER INFORMATION: FGFR1 Alpha Exon
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 57
; OTHER INFORMATION: "IDENTITY OF
; OTHER INFORMATION: NUCLEOTIDE PROVISIONAL"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 117
; OTHER INFORMATION: "IDENTITY OF
; OTHER INFORMATION: NUCLEOTIDE PROVISIONAL"
US-08-371-001-14

Alignment Scores:
Pred. No.: 2,34e-10 Length: 2733
Score: 107.00 Matches: 22
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 1 Gaps: 0

US-09-757-415A-3 (1-22) x US-08-371-001-14 (1-2733)

Qy 1 HisSerGlnMetAlaValHisLysLeuAlaLysSerIleProLeuArgGlnValThr 20
Db 1342 CACAGCCAGATGGCTGTGCACCAAGCTGGCCCAAGAGCATCCCTCTGGCGACAGGTAACA 1401

Qy 21 ValSer 22
Db 1402 GTGTCT 1407

RESULT 8

PCT-US96-00331-14
; Sequence 14, Application PC/TUS9600331
; GENERAL INFORMATION:
; APPLICANT: GENTA INCORPORATED
; TITLE OF INVENTION: METHODS AND COMPOSITION FOR
; TREATING TUMOR CELLS
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; STREET: Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/00331
; FILING DATE: 10 JANUARY 1996
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/371,001
; FILING DATE: 10 JANUARY 1995
; ATTORNEY/AGENT INFORMATION:
; NAME: BIGGS, SUZANNE L.
; REGISTRATION NUMBER: 30,158
; REFERENCE/DOCKET NUMBER: 218/068-PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2733 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 210..467
; OTHER INFORMATION: FGFR1 Alpha Exon
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 57
; OTHER INFORMATION: "IDENTITY OF NUCLEOTIDE PROVISIONAL"

Alignment Scores:
Pred. No.: 2,34e-10 Length: 2733
Score: 107.00 Matches: 22
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 5 Gaps: 0
US-09-757-415A-3 (1-22) x PCT-US96-00331-14 (1-2733)
QY 1 HisSerGlnMetAlaValHisLysLeuAlaLysSerIleProLeuArgGlnValThr 20
DB 1342 CACGCCAGATGGCTGTGCACAAAGCATCCCTCTCGCAGACAGGTAACA 1401

QY

21 ValSer 22

|||||
DB 1402 GTGTCT 1407

RESULT 9

US-07-631-717A-1
; Sequence 1, Application US/07631717A
; Patent No. 5270197
; GENERAL INFORMATION:
; APPLICANT: Yayon, Avner
; APPLICANT: Ornitz, David M.
; APPLICANT: Klagsbrun, Michael
; APPLICANT: Leder, Philip
; TITLE OF INVENTION: SYSTEM FOR ASSAYING BINDING
; TITLE OF INVENTION: TO A HEPARIN-BINDING GROWTH
; TITLE OF INVENTION: FACTOR RECEPTOR
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 50Z or 55SX
; OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)
; SOFTWARE: WordPerfect (Version 5.0)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/631.717A
; FILING DATE: 19901220
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Paul T. Clark
; REGISTRATION NUMBER: 30,162
; REFERENCE/DOCKET NUMBER: 00383/018001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3503
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-07-631-717A-1

Alignment Scores:
Pred. No.: 3,26e-10 Length: 3503
Score: 107.00 Matches: 22
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 1 Gaps: 0

US-09-757-415A-3 (1-22) x US-07-631-717A-1 (1-3503)

QY 1 HisSerGlnMetAlaValHisLysLeuAlaLysSerIleProLeuArgGlnValThr 20

DB 1500 CATAGCCAGATGGCTGTGCACAAAGCTGGCCAAAGCATCCCTCTCGCAGACAGGTAACA 1559

QY 21 ValSer 22

|||||
DB 1560 GTGTCA 1565

RESULT 10

US-08-166-717D-1
; Sequence 1, Application US/08166717D

Patent No. 5789182
GENERAL INFORMATION:
APPLICANT: Yayon, Avner
APPLICANT: Ornitz, David M.
APPLICANT: Klagsbrun, Michael
APPLICANT: Leder, Philip
TITLE OF INVENTION: SYSTEM FOR ASSAYING BINDING
TITLE OF INVENTION: TO A HEPARIN-BINDING GROWTH
TITLE OF INVENTION: FACTOR RECEPTOR
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Clark & Eibing LLP
STREET: 176 Federal Street
CITY: Boston
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02110
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
COMPUTER: IBM COMPATIBLE
OPERATING SYSTEM: WINDOWS 95
SOFTWARE: Wordperfect (Version 7.0)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/166,717D
FILING DATE: 12/14/93
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/631,717
FILING DATE: 12/20/90
ATTORNEY/AGENT INFORMATION:
NAME: Kristina Bieker-Brady
REGISTRATION NUMBER: 39,109
REFERENCE/DOCKET NUMBER: 00383/017002
TELEPHONE: (617) 723-4123
TELEFAX: (617) 723-8962
TELEX:
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 3503
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-166-717D-1
Alignment Scores:
Pred. No.: 3,266-10 Length: 3503
Score: 107.00 Matches: 22
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 1 Gaps: 0
US-09-757-415A-3 (1-22) x US-08-166-717D-1 (1-3503)
QY 1 HisSerGlnMetAlaValHisLysLeuAlaLysSerIleProLeuArgGlnValThr 20
Db 1500 CATAGCCAGATGGCTGTGCACAGCTGGCCAGAGCATCCCTCTGCGCAGACAGGTAACA 1559
QY 21 ValSer 22
Db 1560 GTGTCA 1565
RESULT 11
US-08-471-570-5
Sequence 5, Application US/08471570
Patent No. 5750371
GENERAL INFORMATION:
APPLICANT: IGARASHI, Koichi
APPLICANT: SENOO, Masaharu
APPLICANT: WATANABE, Tatsuya
TITLE OF INVENTION: PROTEIN, DNA AND USE THEREOF
NUMBER OF SEQUENCES: 18

CORRESPONDENCE ADDRESS:
ADDRESSEE: DAVID G. CONLIN; DIKE, BRONSTEIN, ROBERTS &
ADDRESSEE: CUSHMAN
STREET: 130 Water Street
CITY: Boston
STATE: Massachusetts
COUNTRY: US
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/471,570
FILING DATE: 06-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/149,664
FILING DATE:
APPLICATION NUMBER: US 07/743369
FILING DATE: 16-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: LINEK, Ernest V
REGISTRATION NUMBER: 29822
REFERENCE/DOCKET NUMBER: 40897
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 523-3400
TELEFAX: (617) 523-6440
TELEX: 200291 STRE UR
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 1954 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 25..1953
US-08-471-570-5
Alignment Scores:
Pred. No.: 5,578-06 Length: 1954
Score: 83.00 Matches: 18
Percent Similarity: 85.71% Conservative: 0
Best Local Similarity: 85.71% Mismatches: 3
Query Match: 77.57% Indels: 0
DB: 1 Gaps: 0
US-09-757-415A-3 (1-22) x US-08-471-570-5 (1-1954)
QY 2 SerGlnMetAlaValHisLysLeuAlaLysSerIleProLeuArgGlnValThrVal 21
Db 1258 ACCGAGCGGCTGTGCACAGCTGACCAAGCTATCCCCCTCGGAGACAGGTACAGTT 1317
QY 22 Ser 22
Db 1318 TCG 1320
RESULT 12
US-08-471-570-7
Sequence 7, Application US/08471570
Patent No. 5750371
GENERAL INFORMATION:
APPLICANT: IGARASHI, Koichi
APPLICANT: SENOO, Masaharu
APPLICANT: WATANABE, Tatsuya
TITLE OF INVENTION: PROTEIN, DNA AND USE THEREOF
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: DAVID G. CONLIN; DIKE, BRONSTEIN, ROBERTS &
ADDRESSEE: CUSHMAN

STREET: 130 Water Street
CITY: Boston
STATE: Massachusetts
COUNTRY: US
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/471.570
FILING DATE: 06-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/149,664
FILING DATE:
APPLICATION NUMBER: US 07/743369
FILING DATE: 16-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: LINEK, Ernest V
REGISTRATION NUMBER: 29822
REFERENCE/DOCKET NUMBER: 40897
TELEPHONE: (617)523-3400
TELEFAX: (617)523-6440
TELEX: 200291 STRE UR
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 2676 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 25...2331
US-09-471-570-7

Alignment Scores:
Pred. No.: 8.49e-06 Length: 2676
Score: 83.00 Matches: 18
Percent Similarity: 85.71% Conservative: 0
Best Local Similarity: 85.71% Mismatches: 3
Query Match: 77.57% Indels: 0
DB: 1 Gaps: 0

US-09-757-415A-3 (1-22) x US-08-471-570-7 (1-2676)
QY 2 SerGlnMetAlaValHisLysLeuAlaLysSerIleProLeuArgGlnValThrVal 21
Db 1258 AGCCAGCGGCTGTGCACAAAGCTGACCAAAAGTATCCCTCGGGAGACAGGTAAACAGTT 1317
QY 22 Ser 22
Db 1318 TCG 1320

RESULT 13
US-08-451-822A-15
Sequence 15, Application US/08451822A
Patent No. 586388
GENERAL INFORMATION:
APPLICANT: Dionne, Craig A
APPLICANT: Crumley, Greg
APPLICANT: Jaye, Michael C
APPLICANT: Schllessinger, Joseph
TITLE OF INVENTION: Fibroblast Growth Factor Receptors
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: Rhone-Poulenc Rorer Legal Department
STREET: 500 Arcola Road
CITY: Collegeville
STATE: PA

COUNTRY: USA
ZIP: 19426
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/451,822A
FILING DATE: 26-MAY-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/323,430
FILING DATE: 14-OCT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/934,372
FILING DATE: 21-AUG-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/549,587
FILING DATE: 06-JUL-1990
ATTORNEY/AGENT INFORMATION:
NAME: Savitzky, Martin
REGISTRATION NUMBER: 29,699
REFERENCE/DOCKET NUMBER: A0496E
TELEPHONE: (610) 454-3816
TELEFAX: (610) 454-3808
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 3416 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-08-451-822A-15

Alignment Scores:
Pred. No.: 1.18e-05 Length: 3416
Score: 83.00 Matches: 18
Percent Similarity: 85.71% Conservative: 0
Best Local Similarity: 85.71% Mismatches: 3
Query Match: 77.57% Indels: 0
DB: 2 Gaps: 0

US-09-757-415A-3 (1-22) x US-08-451-822A-15 (1-3416)
QY 2 SerGlnMetAlaValHisLysLeuAlaLysSerIleProLeuArgGlnValThrVal 21
Db 1411 AGCCAGCGGCTGTGCACAAAGCTGACCAAAAGTATCCCTCGGGAGACAGGTAAACAGTT 1470
QY 22 Ser 22
Db 1471 TCG 1473

RESULT 14
US-08-323-430-15
Sequence 15, Application US/08323430
Patent No. 6344546
GENERAL INFORMATION:
APPLICANT: Dionne, Craig A
APPLICANT: Crumley, Greg
APPLICANT: Jaye, Michael C
APPLICANT: Schllessinger, Joseph
TITLE OF INVENTION: Fibroblast Growth Factor Receptors
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: Rhone-Poulenc Rorer Legal Department
STREET: 500 Arcola Road
CITY: Collegeville
STATE: PA
COUNTRY: USA
ZIP: 19426
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/323,430
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US to be assigned
FILING DATE: 21-AUG-1992
APPLICATION NUMBER: US 07/549,587
FILING DATE: 06-JUL-1990
ATTORNEY/AGENT INFORMATION:
NAME: Goodman, Rosanne
REGISTRATION NUMBER: 32,534
REFERENCE/DOCKET NUMBER: A0496
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 454-3817
TELEFAX: (215) 454-3808
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 3416 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-323-430-15

Alignment Scores:
Pred. No.: 1.18e-05 Length: 3416
Score: 83.00 Matches: 18
Percent Similarity: 85.71% Conservative: 0
Best Local Similarity: 85.71% Mismatches: 3
Query Match: 77.57% Indels: 0
DB: 4 Gaps: 0

US-09-757-415A-3 (1-22) x US-08-323-430-15 (1-3416)

QY 2 SerGlnMetAlaValHisLysLeuAlaLysSerIleProLeuArgGlnValThrVal 21
Db 1411 AGCCAGCGCGTGTGCACAAAGTATCCCTCGGACAGGTAACAGTT 1470

QY 22 Ser 22
Db 1471 TCG 1473

RESULT 15
US-08-070-165F-5
Sequence 5, Application US/08070165F
Patent No. 5750365
GENERAL INFORMATION:
APPLICANT: Chiu, Ing-Ming
TITLE OF INVENTION: Acidic Fibroblast Growth Factor (aFGF)
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Ing-Ming Chiu
STREET: S2052 Davis Medical Research Center, 480 West
STREET: 9th Avenue
CITY: Columbus
STATE: Ohio
COUNTRY: USA
ZIP: 43210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/070,165F
FILING DATE:
CLASSIFICATION: 435

TELECOMMUNICATION INFORMATION:
TELEPHONE: (614)-293-8093
TELEFAX: (614)-293-5631
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 2675 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: No. 5750365ophthalmus viridescens
DEVELOPMENTAL STAGE: Adult
TISSUE TYPE: Regenerating forelimb blastema
CELL TYPE: Mesenchyme and Epithelium
IMMEDIATE SOURCE:
LIBRARY: lambda gt11
CLONE: KP23-1
POSITION IN GENOME:
UNITS: bp
FEATURE:
NAME/KEY: 5'UTR
LOCATION: 1.324
FEATURE:
NAME/KEY: CDS
LOCATION: 325...2511
FEATURE:
NAME/KEY: 3'UTR
LOCATION: 2512...2675
PUBLICATION INFORMATION:
AUTHORS: Poulin, Matthew L
TITLE: Nucleotide sequences of two new
TITLE: (No. 5750365ophthalmus viridescens) fibroblast growth
TITLE: factor receptor-2 variants
JOURNAL: Biochim. Biophys. Acta
VOLUME: 1220
PAGES: 209-211
DATE: 1994
RELEVANT RESIDUES IN SEQ ID NO: 5: FROM 1 TO 2675
US-08-070-165F-5

Alignment Scores:
Pred. No.: 2.04e-05 Length: 2675
Score: 81.00 Matches: 17
Percent Similarity: 90.48% Conservative: 2
Best Local Similarity: 80.95% Mismatches: 2
Query Match: 75.70% Indels: 0
DB: 1 Gaps: 0

US-09-757-415A-3 (1-22) x US-08-070-165F-5 (1-2675)

QY 2 SerGlnMetAlaValHisLysLeuAlaLysSerIleProLeuArgGlnValThrVal 21
Db 1279 AGCCACCGCTGTGCACAAAGTGTGAGCAAGAGTCTCCCCCTGCGCAGACAGTAACAGTG 1338

QY 22 Ser 22
Db 1339 TCT 1341

Search completed: August 9, 2004, 00:57:59
Job time : 23.7895 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: August 8, 2004, 23:07:35 ; Search time 98.5658 Seconds

(without alignments)
1094.387 Million cell updates/sec

Title: US-09-757-415A-3

Perfect score: 107
Sequence: 1 HSMVAVHKLAKSLRRQVTVS 22

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Zgapop 6.0 , Zgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3222919 seqs, 2451570024 residues

Total number of hits satisfying chosen parameters: 6445838

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US09757415/runat_02082004_101234_6802/app_query.fasta_1.526
-DB=Published Applications NA -QFMT=fastap -SURFIX=p2n.rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09757415 @CGN 1 1 630 @runat_02082004_101234_6802
-NCPUG=6 -ICPU=3 -NO_WMAP -LARGESOURCE -NEG_SCORES=0 -WAIT -DSPELLOCK=100
-LONGFUG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications NA:
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:
13: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:
14: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:
15: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:
16: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:
17: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:
18: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:
19: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
------------	-------	-------	--------	----	-------------

RESULT 1

US-10-641-643-1288
; Sequence 1288, Application US/10641643
; Publication No. US20040077003A1
; GENERAL INFORMATION:
; APPLICANT: Cocks, Benjamin G.
; Susan G. Stuart
; Jeffrey J. Seilhamer

TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL
GENE EXPRESSION

NUMBER OF SEQUENCES: 1508

CORRESPONDENCE ADDRESS:

ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
STREET: 3174 PORTER DRIVE
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2

ALIGNMENTS

1	107	100.0	2360	17	US-10-641-643-1288	Sequence 1288, Ap
2	107	100.0	2470	13	US-10-307-817-5	Sequence 5, Appli
3	107	100.0	2662	15	US-10-204-041-1	Sequence 1, Appli
4	107	100.0	2731	13	US-10-307-817-7	Sequence 7, Appli
5	107	100.0	2737	13	US-10-307-817-3	Sequence 3, Appli
6	107	100.0	3328	15	US-10-101-510-299	Sequence 299, App
7	107	100.0	3328	17	US-10-283-375A-98	Sequence 98, Appl
8	107	100.0	3343	17	US-10-302-812-41	Sequence 41, Appl
9	107	100.0	3454	14	US-10-044-090-48	Sequence 48, Appl
10	107	100.0	3726	9	US-09-925-302-271	Sequence 271, App
11	107	100.0	3726	13	US-09-925-302-271	Sequence 271, App
12	107	100.0	4066	15	US-10-007-926A-182	Sequence 182, App
13	107	100.0	4066	16	US-10-159-563-305	Sequence 305, App
14	107	100.0	4628	13	US-09-805-020-11	Sequence 11, Appl
15	101	94.4	1726	13	US-10-087-192-380	Sequence 380, App
16	101	94.4	3944	13	US-10-087-192-383	Sequence 383, App
17	101	94.4	11294	10	US-09-764-891-7594	Sequence 7594, Ap
18	101	94.4	15123	10	US-09-764-891-7595	Sequence 7595, Ap
19	101	94.4	58215	13	US-10-087-192-379	Sequence 379, App
20	101	94.4	75853	13	US-10-087-192-382	Sequence 382, App
21	87	81.3	505	10	US-09-918-995-4697	Sequence 4697, Ap
22	84	78.5	1451	9	US-09-917-800A-1466	Sequence 1466, Ap
23	83	77.6	2650	10	US-09-954-556-28	Sequence 28, Appl
24	83	77.6	2826	10	US-09-954-556-21	Sequence 21, Appl
25	83	77.6	2868	10	US-09-954-556-19	Sequence 19, Appl
26	83	77.6	2923	10	US-09-954-556-20	Sequence 20, Appl
27	83	77.6	2941	10	US-09-954-556-18	Sequence 18, Appl
28	83	77.6	3080	10	US-09-954-556-25	Sequence 25, Appl
29	83	77.6	3248	17	US-10-302-812-43	Sequence 43, Appl
30	83	77.6	3306	10	US-09-954-556-10	Sequence 10, Appl
31	83	77.6	4268	9	US-09-954-456-293	Sequence 293, App
32	83	77.6	4268	9	US-09-954-456-1599	Sequence 1599, Ap
33	83	77.6	4268	10	US-09-954-556-3	Sequence 3, Appli
34	83	77.6	4268	12	US-09-968-007A-461	Sequence 461, App
35	83	77.6	4574	15	US-10-648-593-132	Sequence 132, App
36	80	74.8	99	15	US-10-199-820-78	Sequence 78, Appl
37	77	72.0	1978	13	US-10-087-192-1130	Sequence 1130, Ap
38	77	72.0	2079	13	US-10-087-192-1133	Sequence 1133, Ap
39	77	72.0	57561	13	US-10-087-192-1129	Sequence 1129, Ap
40	77	72.0	126266	13	US-10-087-192-1132	Sequence 1132, Ap
41	77	72.0	132762	10	US-09-954-556-17	Sequence 17, Appl
42	71	66.4	3025	15	US-09-954-556-23	Sequence 23, Appl
43	71	66.4	4667	15	US-10-007-926A-15	Sequence 15, Appl
44	50	46.7	534	13	US-10-282-122A-21359	Sequence 21359, A
45	50	46.7	711	10	US-09-948-820-36	Sequence 36, Appl

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/641,643
FILING DATE: 14-Aug-2003
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: <Unknown>
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Zeller, Karen J.
REGISTRATION NUMBER: 37,071
REFERENCE/DOCKET NUMBER: PA-0001 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 855-0555
TELEFAX: (650) 845-4166
INFORMATION FOR SEQ ID NO: 1288:
SEQUENCE CHARACTERISTICS:
LENGTH: 2360 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GENBANK
CLONE: g31386
SEQUENCE DESCRIPTION: SEQ ID NO: 1288 :

US-10-641-643-1288

Alignment Scores:
Pred. No.: 8.63e-10 Length: 2360
Score: 107.00 Matches: 22
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 17 Gaps: 0

US-09-757-415A-3 (1-22) x US-10-641-643-1288 (1-2360)

Qy 1 HisSerGlnMetAlaValHisLysLeuAlaLysSerIleProLeuArgGlnValThr 20
Db 1114 CACAGCCAGATGGCTGTGCACAAAGCTGCCAAGAGCATCCCTCTGCGCAGACAGGTAACA 1173
Qy 21 ValSer 22
Db 1174 GTGTCT 1179

RESULT 2

US-10-307-817-5
Sequence 5, Application US/10307817
Publication No. US20040058338A1
GENERAL INFORMATION:
APPLICANT: Agee et al.
TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
FILE REFERENCE: 21402-502C
CURRENT APPLICATION NUMBER: US/10/307,817
CURRENT FILING DATE: 2002-12-02
NUMBER OF SEQ ID NOS: 682
SOFTWARE: CuraSeqList version 0.1
SEQ ID NO 5
LENGTH: 2470
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (116)..(2320)
US-10-307-817-5

Alignment Scores:
Pred. No.: 9.14e-10 Length: 2470
Score: 107.00 Matches: 22
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 13 Gaps: 0

US-09-757-415A-3 (1-22) x US-10-307-817-5 (1-2470)

Qy 1 HisSerGlnMetAlaValHisLysLeuAlaLysSerIleProLeuArgGlnValThr 20
Db 1079 CACAGCCAGATGGCTGTGCACAAAGCTGCCAAGAGCATCCCTCTGCGCAGACAGGTAACA 1138
Qy 21 ValSer 22
Db 1139 GTGTCT 1144

RESULT 3

US-10-204-041-1
Sequence 1, Application US/10204041
Publication No. US20030176443A1
GENERAL INFORMATION:
APPLICANT: STEIN-GERLACH, MATTHIAS
APPLICANT: SALASSIDIS, KONSTADINOS
APPLICANT: BACHER, GERALD
APPLICANT: MULLER, STEFAN
TITLE OF INVENTION: Pyridylpyrimidine Derivatives as Effective Compounds Against Prion
TITLE OF INVENTION: Infections and Prion Diseases
FILE REFERENCE: AXM-007.1P US
CURRENT APPLICATION NUMBER: US/10/204,041
CURRENT FILING DATE: 2002-08-16
PRIOR APPLICATION NUMBER: EP 01111858.5
PRIOR FILING DATE: 2001-05-16
PRIOR APPLICATION NUMBER: PCT/EP02/05420
PRIOR FILING DATE: 2002-05-16
NUMBER OF SEQ ID NOS: 20
SOFTWARE: PatentIn version 3.1
SEQ ID NO 1
LENGTH: 2662
TYPE: DNA
ORGANISM: Homo sapiens
US-10-204-041-1

Alignment Scores:

Pred. No.: 1.01e-09 Length: 2662
Score: 107.00 Matches: 22
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 15 Gaps: 0

US-09-757-415A-3 (1-22) x US-10-204-041-1 (1-2662)

Qy 1 HisSerGlnMetAlaValHisLysLeuAlaLysSerIleProLeuArgGlnValThr 20
Db 1312 CACAGCCAGATGGCTGTGCACAAAGCTGCCAAGAGCATCCCTCTGCGCAGACAGGTAACA 1371
Qy 21 ValSer 22
Db 1372 GTGTCT 1377

RESULT 4

US-10-307-817-7
Sequence 7, Application US/10307817
Publication No. US20040058338A1
GENERAL INFORMATION:
APPLICANT: Agee et al.
TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
FILE REFERENCE: 21402-502C
CURRENT APPLICATION NUMBER: US/10/307,817
CURRENT FILING DATE: 2002-12-02
NUMBER OF SEQ ID NOS: 682
SOFTWARE: CuraSeqList version 0.1
SEQ ID NO 7
LENGTH: 2731
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (116)..(2581)
US-10-307-817-7

```
US-10-307-817-7
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 299
; LENGTH: 3328
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-101-510-299
Alignment Scores:
Pred. No.: 1.04e-09 Length: 2731
Score: 107.00 Matches: 22
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 13 Gaps: 0
US-09-757-415A-3 (1-22) x US-10-307-817-7 (1-2731)
Qy 1 HisSerGlnMetAlaValHisLysLeuAlaLysSerIleProLeuArgGlnValThr 20
Db 1340 CACAGCCAGATGGCTGTGCACAGCTGCCACAGAGCATCCCTCTGCGCAGACAGGTAACA 1399
Qy 21 ValSer 22
Db 1400 GTGTCT 1405
RESULT 5
US-10-307-817-3
; Sequence 3, Application US/10307817
; Publication No. US20040058338A1
; GENERAL INFORMATION:
; APPLICANT: Agee et al.
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 21402-502C
; CURRENT APPLICATION NUMBER: US/10/307,817
; CURRENT FILING DATE: 2002-12-02
; NUMBER OF SEQ ID NOS: 682
; SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 3
; LENGTH: 2737
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (116)..(2587)
US-10-307-817-3
Alignment Scores:
Pred. No.: 1.04e-09 Length: 2737
Score: 107.00 Matches: 22
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 13 Gaps: 0
US-09-757-415A-3 (1-22) x US-10-307-817-3 (1-2737)
Qy 1 HisSerGlnMetAlaValHisLysLeuAlaLysSerIleProLeuArgGlnValThr 20
Db 1346 CACAGCCAGATGGCTGTGCACAGCTGCCACAGAGCATCCCTCTGCGCAGACAGGTAACA 1405
Qy 21 ValSer 22
Db 1406 GTGTCT 1411
RESULT 6
US-10-101-510-299
; Sequence 299, Application US/10101510
; Publication No. US20030148295A1
; GENERAL INFORMATION:
; APPLICANT: WAN, JACKSON
; APPLICANT: WANG, YIXIN
; TITLE OF INVENTION: EXPRESSION PROFILES AND METHODS OF USE
; FILE REFERENCE: 15117.0012
; CURRENT APPLICATION NUMBER: US/10/101,510
; CURRENT FILING DATE: 2002-03-20
; PRIOR APPLICATION NUMBER: 60/276,947
; PRIOR FILING DATE: 2001-03-20
; NUMBER OF SEQ ID NOS: 805
US-10-307-817-7
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 299
; LENGTH: 3328
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-101-510-299
Alignment Scores:
Pred. No.: 1.34e-09 Length: 3328
Score: 107.00 Matches: 22
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 15 Gaps: 0
US-09-757-415A-3 (1-22) x US-10-101-510-299 (1-3328)
Qy 1 HisSerGlnMetAlaValHisLysLeuAlaLysSerIleProLeuArgGlnValThr 20
Db 1180 CACAGCCAGATGGCTGTGCACAGCTGCCACAGAGCATCCCTCTGCGCAGACAGGTAACA 1239
Qy 21 ValSer 22
Db 1240 GTGTCT 1245
RESULT 7
US-10-283-975A-98
; Sequence 98, Application US/10283975A
; Publication No. US20040110792A1
; GENERAL INFORMATION:
; APPLICANT: Ortho-Clinical Diagnostics, Inc.
; TITLE OF INVENTION: Methods For Assessing and Treating Leukemia
; FILE REFERENCE: CDS 293 PCT
; CURRENT APPLICATION NUMBER: US/10/283,975A
; CURRENT FILING DATE: 2002-10-30
; PRIOR APPLICATION NUMBER: 60/340,938
; PRIOR FILING DATE: 2001-10-30
; PRIOR APPLICATION NUMBER: 60/338,997
; PRIOR FILING DATE: 2001-10-30
; PRIOR APPLICATION NUMBER: 60/340,081
; PRIOR FILING DATE: 2001-10-30
; PRIOR APPLICATION NUMBER: 60/341,012
; PRIOR FILING DATE: 2001-10-30
; NUMBER OF SEQ ID NOS: 900
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 98
; LENGTH: 3328
; TYPE: DNA
; ORGANISM: HUMAN
US-10-283-975A-98
Alignment Scores:
Pred. No.: 1.34e-09 Length: 3328
Score: 107.00 Matches: 22
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 17 Gaps: 0
US-09-757-415A-3 (1-22) x US-10-283-975A-98 (1-3328)
Qy 1 HisSerGlnMetAlaValHisLysLeuAlaLysSerIleProLeuArgGlnValThr 20
Db 1180 CACAGCCAGATGGCTGTGCACAGCTGCCACAGAGCATCCCTCTGCGCAGACAGGTAACA 1239
Qy 21 ValSer 22
Db 1240 GTGTCT 1245
RESULT 8
US-10-302-812-41
; Sequence 41, Application US/10302812
; Publication No. US20040087016A1
```

```
; GENERAL INFORMATION:
; APPLICANT: Keating et al.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR CELL DEDIFFERENTIATION AND
; TITLE OF INVENTION: TISSUE REGENERATION
; FILE REFERENCE: HYDR-P02-004
; CURRENT APPLICATION NUMBER: US/10/302,812
; CURRENT FILING DATE: 2002-11-21
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 41
; TYPE: DNA
; LENGTH: 3343
; ORGANISM: Homo sapiens
US-10-302-812-41

Alignment Scores:
Pred. No.: 1.35e-09 Length: 3343
Score: 107.00 Matches: 22
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 17 Gaps: 0

US-09-757-415A-3 (1-22) x US-10-302-812-41 (1-3343)

QY 1 HisSerGlnMetAlaValHisLysLeuAlaLysSerIleProLeuArgArgGlnValThr 20
DB 1228 CACAGCCAGATGGCTGTGCACAAAGCTGCCAAGAGCATCCCTCTGCGCAGACAGGTAACA 1287
QY 21 ValSer 22
DB 1288 GTGTCT 1293

RESULT 9
US-10-044-090-48
; Sequence 48, Application US/10044090
; Publication No. US20020137081A1
; GENERAL INFORMATION:
; APPLICANT: Olga Bandman
; TITLE OF INVENTION: GENES DIFFERENTIALLY EXPRESSED IN VASCULAR TISSUE ACTIVATION
; FILE REFERENCE: PA-0028 US
; CURRENT APPLICATION NUMBER: US/10/044,090
; CURRENT FILING DATE: 2002-01-09
; NUMBER OF SEQ ID NOS: 850
; SOFTWARE: PERL Program
; SEQ ID NO 48
; LENGTH: 3454
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20020137081A1 493848CB1
; LOCATION: 2, 9, 16, 39, 42, 68, 143, 268
; OTHER INFORMATION: a, t, c, g, or other
US-10-044-090-48

Alignment Scores:
Pred. No.: 1.4e-09 Length: 3454
Score: 107.00 Matches: 22
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 14 Gaps: 0

US-09-757-415A-3 (1-22) x US-10-044-090-48 (1-3454)

QY 1 HisSerGlnMetAlaValHisLysLeuAlaLysSerIleProLeuArgArgGlnValThr 20
DB 1078 CACAGCCAGATGGCTGTGCACAAAGCTGCCAAGAGCATCCCTCTGCGCAGACAGGTAACA 1137
QY 21 ValSer 22
DB 1288 GTGTCT 1293
```

```
DB 1138 GTGTCT 1143

RESULT 10
US-09-925-302-271
; Sequence 271, Application US/09925302
; Patent No. US20020044941A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA104
; CURRENT APPLICATION NUMBER: US/09/925,302
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05918
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 896
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 271
; LENGTH: 3726
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (2586)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (3523)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (3664)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (3687)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-925-302-271

Alignment Scores:
Pred. No.: 1.55e-09 Length: 3726
Score: 107.00 Matches: 22
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-09-757-415A-3 (1-22) x US-09-925-302-271 (1-3726)

QY 1 HisSerGlnMetAlaValHisLysLeuAlaLysSerIleProLeuArgArgGlnValThr 20
DB 1051 CACAGCCAGATGGCTGTGCACAAAGCTGCCAAGAGCATCCCTCTGCGCAGACAGGTAACA 1110
QY 21 ValSer 22
DB 1111 GTGTCT 1116

RESULT 11
US-09-925-302-271
; Sequence 271, Application US/09925302
; Publication No. US20030064072A9
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA104
; CURRENT APPLICATION NUMBER: US/09/925,302
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05918
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 896
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 271
; LENGTH: 3726
```

```

; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (2586)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (3523)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (3664)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (3687)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-925-302-271

Alignment Scores:
Pred. No.: 1.55e-09 Length: 3726
Score: 107.00 Matches: 22
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 13 Gaps: 0

US-09-757-415A-3 (1-22) x US-09-925-302-271 (1-3726)

QY 1 HisSerGlnMetAlaValHisLysLeuAlaLysSerIleProLeuArgGlnValThr 20
Db 1051 CACAGCCAGATGGCTGTGCACAGCTGCCAAGAGCATCCCTGCGCAGACAGGTAACA 1110
QY 21 ValSer 22
Db 1111 GTGTCT 1116

RESULT 12
US-10-007-926A-182
; Sequence 182, Application US/10007926A
; Publication No. US20030143539A1
; GENERAL INFORMATION:
; APPLICANT: BERTUCCI, FRANCOIS
; APPLICANT: HOULGATTE, REMI
; APPLICANT: BIRNBAUM, DANIEL
; APPLICANT: NGUYEN, CATHERINE
; APPLICANT: VIENS, PATRICE
; APPLICANT: FERT, VINCENT
; TITLE OF INVENTION: GENE EXPRESSION PROFILING OF PRIMARY BREAST CARCINOMAS
; TITLE OF INVENTION: USING ARRAYS OF CANDIDATE GENES
; FILE REFERENCE: 1546-R-00
; CURRENT APPLICATION NUMBER: US/10/007,926A
; CURRENT FILING DATE: 2001-12-07
; PRIOR APPLICATION NUMBER: 60/254,090
; PRIOR FILING DATE: 2000-12-08
; NUMBER OF SEQ ID NOS: 468
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 182
; LENGTH: 4066
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: fibroblast growth factor receptor 1
; OTHER INFORMATION: (fms-related tyrosine kinase 2, pfeiffer syndrome)
US-10-007-926A-182

Alignment Scores:
Pred. No.: 1.73e-09 Length: 4066
Score: 107.00 Matches: 22
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 15 Gaps: 0

```

US-09-757-415A-3 (1-22) x US-10-007-926A-182 (1-4066)

QY 1 HisSerGlnMetAlaValHisLysLeuAlaLysSerIleProLeuArgArgGlnValThr 20
 Db 1951 CACAGCCAGATGCTGTGTCACAGCTGCGCCAGAGCATCCCTCTGCCAGACAGGTAACA 2010

QY 21 ValSer 22
 Db 2011 GTGTCT 2016
 |||||

RESULT 13
 US-10-159-563-305
 : Sequence 305, Application US/10159563
 : Publication No. US20040009154A1
 : GENERAL INFORMATION:
 : APPLICANT: Khan, Javed
 : APPLICANT: Ringmer, Markus
 : APPLICANT: Peterson, Carsten
 : APPLICANT: Meltzer, Paul
 : TITLE OF INVENTION: SELECTIONS OF GENES AND METHODS OF USING THE SAME FOR
 : FILE REFERENCE: 11613.560S11
 : CURRENT APPLICATION NUMBER: US/10/159,563
 : CURRENT FILING DATE: 2002-12-09
 : PRIOR APPLICATION NUMBER: US 10/133,937
 : PRIOR FILING DATE: 2002-04-25
 : NUMBER OF SEQ ID NOS: 444
 : SOFTWARE: PatentIn version 3.1
 : SEQ ID NO 305
 : LENGTH: 4066
 : TYPE: DNA
 : ORGANISM: Homo sapiens
 US-10-159-563-305

Alignment Scores:
 Pred. No.: 1.73e-09 Length: 4066
 Score: 107.00 Matches: 22
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 16 Gaps: 0

US-09-757-415A-3 (1-22) x US-10-159-563-305 (1-4066)

QY 1 HisSerGlnMetAlaValHisLysLeuAlaLysSerIleProLeuArgArgGlnValThr 20
 Db 1951 CACAGCCAGATGCTGTGTCACAGCTGCGCCAGAGCATCCCTCTGCCAGACAGGTAACA 2010

QY 21 ValSer 22
 Db 2011 GTGTCT 2016
 |||||

RESULT 14
 US-09-805-020-11
 : Sequence 11, Application US/09805020
 : Publication No. US20020086384A1
 : GENERAL INFORMATION:
 : APPLICANT: LEVINE, Zurit
 : TITLE OF INVENTION: SPLICE VARIANTS OF ONCOGENES
 : FILE REFERENCE: 2786-0168P
 : CURRENT APPLICATION NUMBER: US/09/805,020
 : CURRENT FILING DATE: 2001-03-13
 : NUMBER OF SEQ ID NOS: 72
 : SOFTWARE: PatentIn Ver. 2.1
 : SEQ ID NO 11
 : LENGTH: 4628
 : TYPE: DNA
 : ORGANISM: Homo sapiens
 : FEATURE:
 : NAME/KEY: misc feature
 : LOCATION: (1)..(4628)
 : OTHER INFORMATION: any n = a,c,g,t any unknown or other
 US-09-805-020-11

```
Alignment Scores:
Pred. No.:          2.04e-09          Length:          4628
Score:              107.00           Matches:          22
Percent Similarity: 100.00%           Conservative:    0
Best Local Similarity: 100.00%          Mismatches:     0
Query Match:        100.00%           Indels:         0
DB:                 13                Gaps:          0

US-09-757-415A-3 (1-22) x US-09-805-020-11 (1-4628)

QY      1 HisSerGlnMetAlaValHisLysLeuAlaLysSerIleProLeuArgArgGlnValThr 20
      |||
Db      1953 CACAGCCAGATGGCTGTGTGCACAAAGTGGCCAAAGAGCATCCCTCTGCGCAGACAGGTAACA 2012

QY      21 ValSer 22
      |||
Db      2013 GTGTCT 2018

RESULT 15
US-10-087-192-380
; Sequence 380, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: CANCER
; FILE REFERENCE: 52945200122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 380
; LENGTH: 1726
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-087-192-380

Alignment Scores:
Pred. No.:          7.95e-09          Length:          1726
Score:              101.00           Matches:          21
Percent Similarity: 95.45%           Conservative:    0
Best Local Similarity: 95.45%          Mismatches:     1
Query Match:        94.39%           Indels:         0
DB:                 13                Gaps:          0

US-09-757-415A-3 (1-22) x US-10-087-192-380 (1-1726)

QY      1 HisSerGlnMetAlaValHisLysLeuAlaLysSerIleProLeuArgArgGlnValThr 20
      |||
Db      1282 CATAGCCAGATGGCTGTGTGCACAAAGTGGCCAAAGAGCATCCCTCTGCGCAGACAGGTAACA 1341

QY      21 ValSer 22
      |||
Db      1342 GAAAGT 1347

Search completed: August 9, 2004, 02:11:17
Job time : 105.566 secs
```